12-2021

Media as Other Information for Fundamental Valuation

Jiajia Fu

The University of Texas Rio Grande Valley

Jingran Zhao

Volume 23, Number 4 – December 2021
China Accounting and Finance Review
中国会计与财务研究
2021年12月 第23卷 第4期

Media as Other Information for Fundamental Valuation*

Jiajia Fu[1] and Jingran Zhao[2]

Received 2nd of March 2020 Accepted 12th of July 2021

Abstract

The media is an important information intermediary. We investigate the informational role of the media by examining whether media content, measured by the sentiment of news articles, contains information about a firm's fundamental value beyond that conveyed in earnings, book value, and analyst forecasts. We show that incorporating media content into Ohlson's (1995) residual income model generally improves its ability to predict future residual income, explain current stock prices, and predict future stock prices. Our results are strengthened when media coverage is higher and when media sentiment is more dispersed.

Keywords: Media Content, Fundamental Valuation, Mispricing

JEL Classification: G12, G14, G32

* We thank The Hong Kong Polytechnic University for its financial support.
[1] Jiajia Fu, School of Accountancy, The University of Texas Rio Grande Valley, USA.
[2] Corresponding author: Jingran Zhao, School of Accounting and Finance, The Hong Kong Polytechnic University, Hong Kong. Email: jingran.zhao@polyu.edu.hk.

I. Introduction

In his famous fundamental valuation paper, Ohlson (1995) connects firm value with earnings, book value, and other information. Empirical studies find that Ohlson's (1995) residual income model (RIM) predicts firm value better than other valuation models, such as the dividend income model and the abnormal earnings growth model (Penman, 2015).[3] Ohlson's model has also been applied widely in the valuation literature (see Frankel and Lee, 1998; Clarkson *et al.*, 2004; Rhodes-Kropf *et al.*, 2005; Dong *et al.*, 2006). Dechow *et al.* (1999) point out that the key to empirically applying Ohlson's model is to incorporate "other information". This other information refers to the information that reflects a firm's fundamental value, which is not reflected in earnings or book value. It is not easy to find an intuitive measure for the other information, so there has been a demand for future research to put a "face" on it (Beaver, 1999). Most empirical studies that implement the RIM mainly rely on using analyst forecasts to measure the other information (e.g. Frankel and Lee, 1998; Dechow *et al.*, 1999; Bryan and Tiras, 2007).

In this study, we propose to use another measure, media content, to measure the other information in Olson's (1995) model. Specifically, we use the average sentiment of firm-specific news articles to measure the other information and examine whether it can improve the ability of Ohlson's model to predict firm value beyond earnings, book value, and analyst forecasts. We assess the forecast accuracy by testing the model's ability to predict future abnormal earnings and future stock returns and explain concurrent stock prices.

The media is an important information intermediary that generates and disseminates a wide range of information to the public. However, the informational role of the media as a source of valuation has not been well explored relative to the roles of other information intermediaries, such as analysts (Bushee *et al.*, 2010). Diverging from prior studies that examine the market response to media coverage or content, we focus on applying a fundamental valuation model to evaluate the informativeness of the media. We examine whether media content contains information about a firm's fundamental value beyond that conveyed in earnings, book value, and analyst forecasts. We are also interested in how different characteristics of the media affect its role in the assessment of firm value.

We conjecture that media content has information about a firm's fundamental value beyond that conveyed in analyst forecasts. Studies have shown that the media provides information beyond that conveyed in analyst forecasts and improves analyst forecast accuracy (Kross *et al.*, 1990; Cao *et al.*, 2020). Bradshaw *et al.* (2021) also find that compared to analyst forecasts, the media reports more qualitative news that contains value-relevant information about firms. Prior studies show that analysts have a selection bias and tend to be optimistic in their forecasts. They seem to follow the old saying "if you don't have anything good to say, don't say anything at all", in that they tend to follow firms with good prospects and are

[3] We use Ohlson (1995) model, Ohlson's model, and RIM interchangeably throughout the paper.

reluctant to issue negative recommendations (McNichols and O'Brien, 1997). In contrast, journalists seek negative news to attract readership. When journalists release negative news, analysts are less likely to incorporate the news into their forecasts due to their incentives to retain a close relationship with companies. Analysts also provide dated and stale information which does not incorporate the latest news related to the firms they follow (Brown, 1991). The topics of media coverage are also more diverse than those of analyst forecasts. Analyst forecasts typically focus on a company's financial performance, while media articles can cover many topics about a firm that are of interest to the audience, such as the personal affairs of the firm's CEO. With the development of social media, there are many more information providers in the media than analysts. The information from social media has been shown to provide information beyond that provided in analyst forecasts (Gu and Kurov, 2020). Thus, we predict that firm-specific business news articles provide information that reflects a firm's intrinsic value beyond that provided by analyst forecasts.

The media has two roles in affecting market price efficiency—information dissemination and information creation (Drake *et al.*, 2014). We focus on the information creation role because we are interested in the additional information content that the media can provide beyond traditional accounting variables and analyst forecasts. In contrast to news flashes, which are created within minutes from press releases and contain virtually no editorial content, full news articles often combine and process information from a variety of sources and provide investors with reporters' views and analysis that go beyond firm disclosures. Therefore, we use full-article news sentiment when measuring other information.

We compare the forecast accuracy of value estimates from three empirical implementations of Ohlson's model that vary in measuring other information: (1) the simple model that ignores other information, (2) the analyst forecast model that incorporates analyst forecasts into other information, and (3) the media model that incorporates media content beyond the analyst forecast model when measuring other information.

We find that the media model significantly outperforms the other two models in predicting future abnormal earnings and explaining concurrent prices in terms of forecast accuracy. This finding suggests that the media contains information that can be used to measure firms' fundamental value beyond that conveyed by earnings, book value, and analyst forecasts. In addition, we examine whether the market corrects the mispricing if the stock price deviates from the price estimate implied by the model. The underlying rationale is that if the estimated stock price reflects the intrinsic value of a firm, then we should expect stock prices to revert to the intrinsic value in the subsequent period. We find that the price deviation in the media model is associated with a stronger reversal in the subsequent year when compared to the reversals associated with the price deviations from the other models. This result implies that incorporating media information into Ohlson's model results in more reliable value estimates.

We also examine the mechanisms through which the media provides value-relevant information beyond that conveyed in earnings, book value, and analyst forecasts. Specifically, we conduct cross-sectional analyses to examine how media characteristics influence the forecast accuracy of the media model relative to the other two models with respect to predicting abnormal earnings and explaining current stock prices. We find that when there is more media coverage of a firm or when media articles' sentiments are more dispersed, the media model provides more accurate estimates on future abnormal earnings and current stock prices when compared to the simple model and the analyst forecast model. This finding is consistent with the conjecture that news articles that are more investigative in nature tend to contain more information beyond that provided by fundamental accounting variables and analyst forecasts.

This study contributes to the literature in several ways. First, it adds to the valuation model literature by incorporating media content when modelling the dynamics of the other information in Ohlson's (1995) model. Prior empirical studies use analyst forecasts to measure the other information (e.g. Frankel and Lee, 1998; Dechow *et al.*, 1999; Bryan and Tiras, 2007). We show that incorporating media information into the other information increases the ability of Ohlson's model to predict valuation estimates. This outcome is also consistent with the conceptual definition of the other information—information that reflects a firm's fundamental value but is not conveyed through earnings and book value.

Second, this study extends the literature that examines the business press as an information intermediary in the capital markets. We provide a different perspective and methodology to test the informational role of the media. Diverging from prior studies that examine the association between the media and investors' responses to accounting information or information asymmetry (Bushee *et al.*, 2010; Drake *et al.*, 2014), we examine whether media information improves the reliability of the intrinsic value estimates derived from Ohlson's model. In addition, while prior research focuses on media coverage (Bushee *et al.*, 2010; Drake *et al.*, 2014; Fang and Peress, 2009), we show that media content (measured by the sentiment of news articles), another key feature of the media, contains information beyond that provided by accounting data and analyst forecasts. We also show that such an effect is more pronounced when media coverage is higher and when media sentiment is more dispersed.

The rest of the paper is organised as follows. Section II reviews the prior literature and presents our hypotheses. Section III describes the data and the implementation of Ohlson's (1995) model in three ways. Section IV presents empirical results and main findings. Section V concludes the paper.

II. Related Literature and Hypothesis Development

2.1 Media's Impact on the Capital Market

The media plays an important and ever increasing role in our daily life and the financial market. As media data become more accessible, research on the media is emerging. Research that examines the impact of the media on the financial capital market focuses on how the market reacts to media coverage (Sant and Zaman, 1996; Barber and Odean, 2008). An early work by Foster (1979) finds that the stock market reacted strongly to articles written by Abraham Briloff, a famous critic of financial reporting standards. The stocks of firms whose accounting practices Briloff criticised dropped by 8% on average on the days his articles were published. Klibanoff *et al.* (1998) show that the response of closed-end country fund prices to country-specific news is greater when the news appears on the front page of the *New York Times*. In a larger scale empirical study, Tetlock (2007) analyses the linguistic content of articles from the *Wall Street Journal* and reports that media pessimism predicts downward pressure and a subsequent reversion to fundamental values. His findings suggest that media content does not contain new information about firms' fundamental values and that the media merely serves as a proxy for investor sentiment.

Recent studies show that the media helps to improve the information environment of the stock market. Tetlock *et al.* (2008) document that the fraction of negative words used in news stories predicts future earnings and stock returns. Even though stock prices briefly underreact to the information embedded in the negative words, investors quickly incorporate the information into stock prices. Fang and Peress (2009) show that stocks with high media coverage earn lower returns, suggesting that information dissemination by the media alleviates informational friction in the stock market, which leads to lower cost of capital. Engelberg and Parsons (2011) find that investors with access to different media attention for the same information event behave differently in trading, suggesting a causal impact of media reporting on stock market reactions to corporate events. Using national newspaper strikes from different countries as exogenous shocks, Peress (2014) shows that media attention improves the dissemination of information among investors and its incorporation into stock prices. Kothari *et al.* (2009) find that positive (negative) press coverage decreases (increases) firms' cost of capital, return volatility, and analyst forecast dispersion. Soltes (2010) documents that firm-initiated disclosures disseminated through the press reduce bid-and-ask spreads, increase trading volume, and lower idiosyncratic volatility. Further, Bushee *et al.* (2010) find that media attention reduces information asymmetry (i.e. lower spreads and greater depth) around earnings announcements. Drake *et al.* (2014) demonstrate that media attention can help the market understand the accounting information contained in financial statements. They find that press coverage of annual earnings announcements can mitigate cash flow mispricing, but not accruals mispricing. The authors conjecture that the complexity of accruals may hinder reporters from understanding and communicating the implications to readers. This result suggests that even though the media has the ability to interpret financial reporting, this ability is limited and incomplete.

Prior literature on the media relies on the stock market reaction to capture the media's information content. Even though the market is highly efficient, there are still times when the stock prices deviate from the firm's fundamental value and investors may not react fully or rationally to the information in the media (Tetlock, 2007; Tetlock *et al.*, 2008). Instead of relying on stock market prices, we use the RIM to evaluate whether the media can be used as a valuable information source for fundamental valuation.

2.2 Ohlson's (1995) Model

Ohlson's (1995) RIM reformulates the dividend discount model and incorporates accounting information into valuation. Penman (2005) calls the residual income valuation model the "centrepiece" of accounting-based valuation. Since its introduction, Ohlson's model has been widely studied and applied by both accounting and finance researchers (e.g. Clarkson *et al.*, 2004). The model's attraction lies in its simplicity of using two accessible accounting measures, book value and earnings, to value a firm. Prior studies find the model fits better than other valuation models, such as the dividend discount model and the abnormal earnings growth model (see Penman, 2005; Lundholm and O'Keefe, 2001). Ohlson's model provides more accurate forecasts of a firm's value than the abnormal earnings growth model (Brief, 2007). Ohlson's model performs even better when the forecast horizon extends from two to five years (Jorgensen *et al.*, 2011). Moreover, the model does well in detecting misvaluation and predicting future stock prices (see Frankel and Lee, 1998; Lee *et al.*, 1999). Many studies apply the model to link mispricing and other corporate activities. For example, Rhodes-Kropf *et al.* (2005) apply the model to detect misvaluation and find a connection between misvaluation and merger activities. Dong *et al.* (2006) also use the model to study investor misvaluation and various aspects of takeover activity.

In addition to its use in empirical studies, the RIM also provides a foundation for theoretical studies. For example, Gode and Ohlson (2004) modify the RIM's constant interest rate assumption and develop a valuation model with changing interest rates. Ohlson and Juettner-Nauroth (2005) relax the RIM's clean surplus assumption and model the stock price as a function of future earnings and dividends. Given its importance in both the finance and accounting literature, we choose to contribute to the implementation of Ohlson's (1995) RIM model by incorporating media content into it.

Besides the contribution Ohlson (1995) brings to the literature, many researchers encounter challenges to implementing the model, and some implement the model incorrectly (Lo and Lys, 2000). Lo and Lys (2000) find that many studies did not implement the information dynamics that are the key feature of the model, which resulted in mixed findings on the effectiveness of the Ohlson (1995) model. The challenge of implementing the information dynamics of the model is to find a good proxy for the "other information". Aside from earnings and book value, which are easy to measure, Ohlson (1995) incorporates the other information, which reflects a firm's fundamental value not reflected in earnings and book value. Scholars mainly rely on using analyst forecasts to measure the other information (e.g. Frankel and Lee, 1998; Dechow *et al.*, 1999). Bryan and Tiras (2007) regress analyst forecasts for year t+1 on current earnings and book value in year t and use the residual from the regression as the proxy for the other information. We evaluate whether media content can be an effective source of other information in Ohlson's (1995) model.

2.3 Hypothesis Development

The business press provides information that captures hard-to-quantify aspects of firm fundamentals (Tetlock *et al.*, 2008). In contrast to financial statements that are issued four times a year at most, business press articles are released to the market on a daily basis. The quantity of information from the media is likely to be substantial because hundreds of articles are released through newspapers, the internet, and social media. Zhao (2017) documents that firms with better coverage of information intermediaries (analysts, institutional investors, shareholders) are associated with more efficient stock prices (i.e. less stock price deviation from the firm's intrinsic value, where the intrinsic value is estimated from Ohlson's model). We extend Zhao (2017) and examine the informational role of the media by examining whether incorporating media information into the information dynamics of Ohlson's model improves the measurement of a firm's intrinsic value. Prior research, as discussed in section 2.1, generally supports the informational role of the media in the capital market. This leads to our first hypothesis:

H1: The media contains incremental information about a firm's value that is not reflected in earnings and book value.

However, this hypothesis has its critics. Unlike financial statements, the business press is not strictly monitored by regulatory bodies, such as the Securities and Exchange Commission, or audited by a third party, such as auditors. Journalists have incentives to generate news articles that are attractive to readers. Prior literature shows that the media creates sensational information to entertain the general public. For example, Jensen (1979) argues that the media provides simple answers to complex problems to stimulate people's curiosity and attention. Prior studies also show that when media outlets face competition from peers, journalists intentionally generate biased news to attract readership (e.g. Baron, 2006; Mullanianthan and Shleifer, 2005). Gentzkow and Shapiro (2010) find that US daily newspapers generate political news to meet the demand from local readers in order to maximise their own profits. DeAngelo *et al.* (1994, 1996) use evidence from the bond market to show that sensational news from the media can lead to suboptimal economic consequences. Core *et al.* (2008) find that the media tends to cover CEOs with more option exercises because these CEO compensation practices are more sensational to readers. Given that, it is unclear whether media sentiment contains incremental information about a firm's fundamental value that is not reflected in earnings and book value.

Our second hypothesis concerns whether the media contains information beyond that conveyed in analyst forecasts to help evaluate a firm's fundamental value. Along with earnings and book value, analyst forecasts are a widely used information source that reflects a firm's fundamental value. The media, on the other hand, is also an important information source for valuations. It covers more broad topics and disseminates timely value-relevant information on firms to both sophisticated and unsophisticated investors as well as other market participants (Bushee *et al.*, 2010). Bushee *et al.* (2010) show that firm-specific media coverage increases market depth and reduces the bid-ask spread around earnings announcements. Engelberg and Parsons (2011) find that the media coverage of earnings announcements in local newspapers stimulates local trading activities.

Moreover, studies have shown that the media provides incremental information to analysts and improves analyst forecast accuracy (Kross *et al.*, 1990; Cao *et al.*, 2020). Bradshaw *et al.* (2021) show that the media reports more qualitative news that contains incremental value-relevant information about a firm. Prior studies also show that analysts have a selection bias and tend to be optimistic in their forecasts. They seem to follow the old saying "if you don't have anything good to say, don't say anything at all", in that they tend to follow firms with good prospects and are reluctant to issue negative recommendations (McNichols and O'Brien, 1997). In contrast, journalists seek negative news to attract readership. When journalists release negative news, analysts are less likely to incorporate the news into their forecasts due to their incentives to retain a close relationship with the companies. Analysts also provide dated and stale information which does not incorporate the latest news related to the firms they follow (Brown, 1991). The topics of media coverage are also more diverse than analyst forecasts. Analyst forecasts typically focus on a company's financial performance, while media articles can cover many topics about the firm that are of interest to their audience, such as the personal affairs of the firm's CEO. With the development of social media, there are many more information providers in the media than analysts. The information from social media has been shown to provide information incremental to analyst forecasts (Gu and Kurov, 2020). If this is the case, then we predict that firm-specific business news articles provide information that reflects a firm's intrinsic value beyond that conveyed in analyst forecasts.

The criticism of this prediction is that analysts are better at acquiring and processing firm information than journalists. Analysts maintain a close relationship with public companies' management so that they can provide superior information and recommendations to their clients. Journalists also tend to lack the financial expertise that analysts possess. Therefore, it is likely that the media provides little information about a firm's intrinsic value beyond that conveyed in analyst forecasts.

A debate exists on whether the media's effect on the pricing dynamics is related to its information dissemination role, its information creation role, or both (Bushee *et al.*, 2010;

Soltes, 2010; Drake *et al.*, 2014). In our setting, both roles are likely to contribute to the information content in the media that is incremental to accounting data and analyst forecasts. The information dissemination role can emphasise the importance of certain news related to a firm. The information creation role of the media can generate information by covering comprehensive news topics about a firm that analysts may fail to process in a timely way or simply deem irrelevant to the firm's value. Thus, we state our second hypothesis in the alternative form as follows:

H2: The media contains incremental information about a firm's value that is not reflected in earnings, book value, and analyst forecasts.

III. Research Design

3.1 Data and Sample

We obtain media data from RavenPack, an aggregator for business press articles. RavenPack gathers and analyses information from three major sources. The first source is Dow Jones, which accesses its information from the Dow Jones newswires, regional editions of the *Wall Street Journal*, and *Barron's*. Its second source is the Web Edition, which collects its information from business publishers, national and local news, blog sites, and government and regulatory updates. The third source is the PR Edition, which collects data from press releases and regulatory, corporate, and news services, including PR Newswire, the CNW Group (formerly the *Canadian News Wire*), and the Regulatory News Service.

We use RavenPack for several reasons. First, it automatically classifies news into firm-specific and performance-related categories by using proprietary text and part-of-speech tagging. It also identifies firms' actions. For example, in a news story with the headline "IBM Completes Acquisition of Telelogic AB", the tag "acquisition acquirer" indicates that IBM is involved in an acquisition event and is the acquiring company and the tag "acquisition-acquiree" indicates that Telelogic is the acquired company. Second, RavenPack classifies news as flashes (i.e. a headline with no body text) or full articles (i.e. at least one paragraph). This classification allows us to differentiate between information dissemination and creation. It is important to be able to differentiate the news flashes from full articles in the current study because our study focuses on the information content that is generated by journalists. The information dissemination of news flashes should not have an impact on the measurement of a firm's fundamental value. Third, RavenPack differentiates between news from the media and news from firms. This feature is also critical to our study because it is important to determine whether the news comes from the firm or the media so that we can examine the incremental power the media has to reflect a firm's value that is not reflected through the firm's financial statement. Ohlson's (1995) model requires other information that reflects firm value that is not already reflected through earnings and book value. Lastly, RavenPack assigns a relevance score ranging from 0 to 100 to each news article, indicating the firm's prominence in the article. For news stories with multiple entities, RavenPack identifies which entity plays the key role.

We obtain financial data from the Compustat annual database, stock return data from CRSP, and analyst forecast data from I/B/E/S summary files. The intersection of the RavenPack data with these additional data sources yields an initial sample from 2000 to 2017.[4] We start with 2000 because RavenPack's news coverage begins in 2000. We then exclude firm-year observations with negative total assets or book value of equity. Sample sizes vary across different test specifications due to data availability and are noted in the tables. Table 1 presents the descriptive statistics for the key variables. The mean sentiment score of media articles is close to zero (mean = -0.005, median = -0.003), indicating that our sample is not biased toward firms with good news or bad news.

Table 1 Descriptive Statistics

Variable	N	Mean	Median	Std Dev	25th Pctl	75th Pctl
x_t^a	24154	-0.041	0.011	0.218	-0.031	0.031
m	23988	-0.005	-0.003	0.023	-0.011	0.005
Media Difference	23046	0.023	0.020	0.021	0.010	0.030
Media Dispersion	22810	0.034	0.024	0.029	0.015	0.042
Media Coverage	23046	0.430	0.280	0.517	0.120	0.560
Hard News Freq	23212	0.216	0.125	0.254	0.000	0.333
$Return_{t+1}$	21338	0.119	0.088	0.412	-0.145	0.321
$AbRet_{t+1}$	21338	0.004	-0.024	0.348	-0.216	0.175
Inst Holding	23212	0.660	0.716	0.283	0.451	0.876
Analyst Coverage	23212	2.204	2.197	0.728	1.792	2.773
#Mgmt Guidance	23212	0.792	0.000	0.957	0.000	1.792
Size	23212	7.270	7.161	1.717	6.024	8.367
BTM	23212	0.487	0.418	0.313	0.260	0.637

This table reports the descriptive statistics for abnormal earnings, media content, and other firm characteristic variables. The sample consists of firm-year observations from 2000 to 2017. All continuous variables are winsorised at the 1st and 99th percentiles and defined as in Appendix A.

Table 2 presents the Pearson correlations of media characteristics and other key firm fundamental variables. We use analyst coverage, institutional holdings, and the frequency of management guidance to proxy for a firm's information environment. *Media Dispersion* and *Media Coverage* are positively correlated with the three proxies of information environment, suggesting that firms with a better information environment tend to have more dispersed media sentiment scores and higher media coverage.

4 As we use one-year ahead earnings and returns when calculating dependent variables, the actual financial data used extend to 2018 and the return data extend to early 2019.

Table 2 Pearson Correlations

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	*Media Difference*	1.000								
(2)	*Media Dispersion*	0.482*	1.000							
(3)	*Media Coverage*	-0.048*	0.048*	1.000						
(4)	*Hard News Freq*	0.099*	0.430*	-0.180*	1.000					
(5)	*Inst Holding*	-0.006	0.033*	0.095*	-0.080*	1.000				
(6)	*Analyst Coverage*	-0.093*	0.147*	0.388*	0.073*	0.270*	1.000			
(7)	*#Mgmt Guidance*	-0.025*	0.062*	0.124*	0.078*	0.154*	0.270*	1.000		
(8)	*Size*	-0.125*	0.131*	0.526*	0.046*	0.187*	0.684*	0.211*	1.000	
(9)	*BTM*	0.040*	0.053*	-0.207*	0.045*	-0.005	-0.224*	-0.089*	-0.355*	1.000

This table reports Pearson correlations of key media characteristics and other firm fundamental variables. All variables are defined as in Appendix A. * shows significance at least at the 5% level.

3.2 Implementation of Ohlson's (1995) Model

In Ohlson's (1995) RIM, stock prices are linked with accounting variables in a linear fashion as follows:

$$P_t = b_t + \sum_{\tau=1}^{\infty} \frac{E_t[x_{t+\tau}^a]}{(1+r)^\tau} \tag{1}$$

($P_t = b_t$ + present value of expected future abnormal earnings)

where P_t is the equity value at *t*; b_t is the book value of equity at *t*; and the abnormal earnings are $x_t^a = x_t - rb_{t-1}$, with *r* being the discount rate.

To measure the present value of expected future abnormal earnings, Ohlson (1995) models the autoregressive process of abnormal earnings and the "other information" about future abnormal earnings that is not reflected in current abnormal earnings; accordingly,

$$x_{t+1}^a = \omega x_t^a + v_t + \varepsilon_{1,t+1} \tag{2}$$

$$v_{t+1} = \gamma v_t + \varepsilon_{2,t+1}. \tag{3}$$

where v_t is other information (i.e. information about future abnormal earnings not in the current abnormal earnings); $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are unpredictable, mean-zero disturbance terms; and ω and γ are fixed-persistence parameters.

Dechow *et al.* (1999) and Ohlson (2001) emphasise the importance of the autoregressive behaviour of residual income and estimate the related "persistence" parameter (i.e. ω and γ). Measuring the persistence parameters allows prices to be depicted as a linear combination of book value, current abnormal earnings, and another information variable. We follow Dechow *et al.* (1999) and measure the autoregressive parameter, ω^c, of conditional abnormal earnings by constructing five variables that they hypothesise to be associated with cross-sectional variation in the persistence of abnormal earnings. We run autoregressive regressions in which each of the five determinants are included as interactive effects, as in the following model:

$$x_t^a = \omega_0 + \omega_1 x_{t-1}^a + \omega_2(x_{t-1}^a q1_{t-1}) + \omega_3(x_{t-1}^a q2_{t-1}) + \omega_4(x_{t-1}^a q3_{t-1}) + \omega_5(x_{t-1}^a q4_{t-1}) + \omega_6(x_{t-1}^a q5_{t-1}) + \varepsilon_t \quad (4)$$

where $q1$ is the absolute value of abnormal earnings for year *t* divided by book value of equity at the beginning of year *t*; $q2$ is the absolute value of special items divided by book value of equity at the beginning of year *t*; $q3$ is the absolute value of accruals divided by book value of equity at the beginning of year *t*; $q4$ is dividends paid during year *t* divided by earnings before extraordinary items for year *t*; and $q5$ is the first-order autoregressive coefficient from an abnormal earnings autoregression for all firms in the same two-digit SIC code as the observation. The autoregression is conducted using all available firms on Compustat in the same two-digit SIC code from 1960 to year *t*.

We then use the parameter estimates from Eq. (4) and the actual firm-year values of the five determinants to compute the estimate for ω^c in the following model:

$$\omega^c = \omega_1 + \omega_2 q1_t + \omega_3 q2_t + \omega_4 q3_t + \omega_5 q4_t + \omega_6 q5_t \quad (5)$$

Using a process similar to our estimation of the conditional value of the autoregressive parameter ω^c of abnormal earnings, we then estimate the conditional value of the autoregressive parameter γ to incorporate media data in the measurement of the "other information". To determine this value, we add the media content variable (m_t) into model (3) because the sentiment of news has been documented to be associated with analyst forecasts, v_t (Bradshaw *et al.*, 2021).

The media variable, m_t, reflects the average opinion of news published about a firm in fiscal year *t*. RavenPack generates a variable called the *Composite Sentiment Score* (CSS), which represents the news sentiment about a firm by combining various textual analysis techniques. It determines the score by using an algorithm that matches stories categorised by financial experts as having short-term positive or negative financial or economic impact. The algorithm also interprets actual figures, estimates, ratings, analyst revisions, and recommendations disclosed in news articles.[5] CSS scores range from 0 to 100, with a score above (below) 50 indicating positive (negative) news and a score equal to 50 indicating natural news. We follow prior research and apply a linear transformation to the CSS, and we define media content m_t to be the average (CSS-50)/100 for news articles released for each firm within each fiscal year. We estimate the regression below:

5 See detailed variable definition in Appendix A. Examples of prior studies that use the CSS score to measure news sentiment include Bushman *et al.* (2017), Chen *et al.* (2021), and von Beschwitz *et al.* (2020).

$$v_t = \gamma_0 + \gamma_1 v_{t-1} + \gamma_2(v_{t-1}m_{t-1}) + \varepsilon_t. \quad (6)$$

We compute the estimate for each firm-year using the parameter estimate from Eq. (6) and the years' actual values of m as follows:

$$\gamma^c \ = \gamma_1 + \gamma_2 m_t. \quad (7)$$

After obtaining both persistence parameters, ω^c and γ^c, we follow Dechow *et al.* (1999) and compare the forecast accuracy of abnormal earnings and stock prices from three empirical implementations of Ohlson's model that vary in measuring other information: (1) the simple model that ignores other information, (2) the analyst forecast model that incorporates analyst forecast into other information, and (3) the media model that incorporates media content beyond the analyst forecast model when measuring other information. We test whether the persistence parameters, ω^c and γ^c, provide the most accurate forecasts for future abnormal earnings and best explain stock prices and predict future returns. Empirical tests are described in detail in section 3.3.

3.3 Empirical Tests

We conduct two main sets of analysis to test the hypotheses. First, we compare the bias and accuracy of the predictions of next period abnormal earnings generated by each of the following models:

(I) Predictions for model ignoring "other information": $E[x^a_{t+1}] = \omega^c x^a_t$

(II) Prediction for model incorporating "other information" using analyst forecasts: $E[x^a_{t+1}] = f^a_t$

(III) Prediction for model incorporating "other information" using analyst forecasts and media content: $E[x^a_{t+1}] = \omega^u x^a_t + \gamma^c v_t$

x^a_{t+1} is abnormal earnings for year t+1, where abnormal earnings is defined as $x^a_t = x_t - r.b_{t-1}$. x_t denotes earnings before extraordinary items and discontinued operations for year t. r denotes the discount rate, and b_t denotes the book value of equity at the end of year t. ω^c is the predicted value of ω from the regression model (4). f^a_t is consensus analyst forecast for year t+1 abnormal earnings, measured as $f_t - r.b_t$, where f_t is consensus analyst forecast for year t+1 earnings in the first month following the earnings announcement of year t. This ensures that all of the forecasting variables are measured at similar points in time. γ^c is calculated as in model (7).

Following Dechow *et al.* (1999), we generate three statistics that evaluate bias and accuracy from each of the three models. Forecast errors (FEs) in earnings models are computed by subtracting the predicted future earnings of year t+1 from the realised earnings of year t+1 scaled by market value at the end of year t. Three metrics include mean FE, mean absolute FE, and mean squared FE.

Next, we evaluate the relative ability of the three models to explain contemporaneous stock prices. The three models are as follows:

(IV) Price estimate for model ignoring “other information”: $P_t = b_t + \frac{\omega^c}{1 + r - \omega^c} x_t^a$

(V) Price estimate for model incorporating “other information” using analyst forecasts:

$$P_t = b_t + \frac{\omega^c}{1 + r - \omega^c} x_t^a + \frac{1 + r}{(1 + r - \omega^u)(1 + r - \gamma^\omega)} v_t$$

(VI) Price estimate for model incorporating “other information” using analyst forecasts and media content: $P_t = b_t + \frac{\omega^c}{1 + r - \omega^c} x_t^a + \frac{1+r}{(1 + r - \omega^u)(1 + r - \gamma^c)} v_t$

b_t is the book value of equity at the end of year t. x_{t+1}^a is abnormal earnings for year t+1, where abnormal earnings is defined as $x_t^a = x_t - r.b_{t-1}$. x_t denotes earnings before extraordinary items and discontinued operations for year t. r denotes the discount rate. ω^c is the predicted value of ω from the regression model (4). γ^ω is the first-order autoregression coefficient for the other information variable, v_t, and is estimated using all historically available data through the forecast year in a pooled time-series cross-sectional regression. v_t is defined as $v_t = f_t^a - \omega^c x_t^a$. f_t^a is the consensus analyst forecast of abnormal earnings for year t+1 measured in the first month following the announcement of earnings for year t. In other words, $f_t^a = f_t - r.b_t$. γ^c is calculated as in model (7).

Using a similar approach as for the abnormal earnings prediction, we compare model performance with three metrics. FEs in price models are computed by subtracting the predicted stock price for year t+1 from the observed stock price at the end of the month following the announcement of earnings for year t, scaled by the observed price. Metrics to evaluate bias and accuracy include mean FE, mean absolute FE, and mean squared FE.

If the media contains incremental information about a firm’s value that is not reflected in earnings, book value, and analyst forecasts, then we should expect that the model that incorporates the other information with analyst forecasts and media content, in both abnormal earnings and price tests, has the smallest mean FE, mean absolute FE, and mean squared FE.

IV. Empirical Results

4.1 Prediction of Next Period Abnormal Earnings

Table 3 reports the metrics on the bias and accuracy of the predictions of next period abnormal earnings generated by each of the valuation models. We use 8% as the approximation of the long-run average realised return on US equities. Theoretically, the discount rate r should be firm specific, reflecting the compensation that equity investors demand for the risk they take to invest in the stock. However, it is difficult to empirically determine the value of r. Because r enters the valuation models in a similar fashion, its variations do not materially affect the model in empirical tests (Dechow *et al.*, 1999).[6]

---

[6] Figure 13 of Damodaran (2020) reports the expected returns on US stocks year by year from 1961 through

FE measures forecast bias, while absolute FE and squared FE measure forecast accuracy. We first report the mean statistics of these metrics for each of the three models, followed by a comparison of the performance of the media model against that of the simple model and the analyst forecast model. The analyst forecast model reports the most negative mean FEs (mean FE = -0.020), consistent with the over-optimism of analysts' forecasts documented in the analyst forecast literature. FE difference is significantly negative in the comparison of the media model versus the simple model but significantly positive in the comparison of the media model versus the analyst forecast model. These outcomes indicate that the predicted abnormal earnings generated from the media model are more biased than those from the simple model but less biased than those from the analyst forecast model. Results on the absolute FE difference and the squared FE difference show that the media model generates lower absolute FE and lower squared FE than the simple model and the analyst forecast model. In other words, the media model produces the smallest magnitude for both accuracy metrics (mean absolute FE = 0.049, mean squared FE = 0.009). In sum, the results in Table 3 provide evidence that the model incorporating media information outperforms the other two models.

Table 3 Earnings Prediction Error Metrics of Three Models

	Mean FE	Mean absolute FE	Mean squared FE
(1) Simple	0.000	0.053	0.011
(2) AF	-0.020	0.056	0.014
(3) Media	-0.009	0.049	0.009

	FE Difference	t-statistic	Absolute FE Difference	t-statistic	Squared FE Difference	t-statistic
(3)-(1)	-0.009***	(-28.42)	-0.004***	(-14.32)	-0.002***	(-14.83)
(3)-(2)	0.010***	(23.73)	-0.007***	(-17.95)	-0.005***	(-21.67)

This table reports and compares error metrics that measure the bias and accuracy of the predictions of next period abnormal earnings generated by each of the earnings valuation models. The sample consists of 24,154 firm-year observations from 2000 to 2017. The forecast error (FE) is computed by subtracting the forecasted abnormal earnings from the actual abnormal earnings for year t+1. T-statistics are reported in brackets. Valuation model variables are defined in Appendix A. *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

4.2 Explanation of Contemporaneous Stock Prices

Next, we evaluate the relative ability of the three models to explain contemporaneous stock prices. The results are reported in Table 4. The FEs in Table 4 are calculated as the difference between the market price and the predicted price, scaled by the market price. As in the earnings model, we use a discount rate of 8%. All three models generate large positive mean FEs, indicating that they largely undervalue stocks relative to the market. The

2018 and shows that the average expected returns from 2000 to 2017 (our sample period) is approximately 8%. Our inferences remain unchanged if using a discount rate from 5% to 10% (untabulated).

undervaluation is smallest for the analyst forecast model (FE = 0.313), which is somewhat consistent with the results in Table 3 showing that analyst forecasts of earnings are the most optimistic ones. Significantly negative statistics on absolute FE difference and squared FE difference indicate that the media model on average presents smaller absolute FE and squared FE than the simple model and the analyst forecast model. This result suggests that the media model outperforms the other two models in terms of the accuracy in explaining contemporaneous stock prices.

Table 4 Price Prediction Error Metrics of Three Models

	Mean FE	Mean absolute FE	Mean squared FE
(1) Simple	0.559	0.579	0.445
(2) AF	0.313	0.647	0.923
(3) Media	0.466	0.531	0.404

	FE Difference	t-statistic	Absolute FE Difference	t-statistic	Squared FE Difference	t-statistic
(3)-(1)	-0.093***	(-39.07)	-0.048***	(-27.52)	-0.035***	(-9.37)
(3)-(2)	0.153***	(33.25)	-0.128***	(-29.83)	-0.667***	(-25.98)

This table reports and compares error metrics that measure the bias and accuracy of the predictions of next period price generated by each of the price valuation models. The sample consists of 23,212 firm-year observations from 2000 to 2017. The forecast error (FE) is computed by subtracting the forecasted abnormal earnings from the actual abnormal earnings for year t+1. T-statistics are reported in brackets. Valuation model variables are defined in Appendix A. *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

4.3 Prediction of Future Stock Returns

So far, our pricing tests compare the ability of the valuation models to explain contemporaneous stock prices. By considering the best valuation model to be the one that best explains the current prices, we assume that the market price reflects the fundamental value of firms (i.e. market efficiency). However, it is possible that stock prices temporarily deviate from their fundamental values. We examine whether investors incorporate information from earnings, book values, analyst forecasts, and media into their investment decisions. Specifically, we test whether observed stock prices revert toward the fundamental values implied by the competing models. If the valuation models reflect the intrinsic value of firms and if investors fail to fully incorporate information in earnings, book values, analyst forecasts, or the media, then we should expect that stock prices will revert to the fundamental values in the future.

We form a zero-investment portfolio based on the deciles of the FEs from the valuation models, measured as the difference between stock prices and the intrinsic model values, divided by stock prices. Higher FEs represent more overpricing of the current stock prices and therefore predict more negative returns in the subsequent year. In other words, lower deciles are stocks that are underpriced relative to fundamental value and are expected to have higher future stock returns. Higher deciles consist of overpriced stocks that are expected to experience lower future returns. The hedged return on the portfolio of the price FEs would indicate the magnitude of price reversion in the subsequent year. The valuation model that best captures the intrinsic firm value is expected to have the strongest price reversion.

Panel A of Table 5 reports the equal-weighted portfolio results. Both buy-and-hold returns and abnormal returns (i.e. DGTW firm characteristic-adjusted returns) are employed to confirm the robustness of the results. DGTW return subtracts, from each stock return, the return on a portfolio of firms matched on market equity, market-book, and prior one-year return quintiles. The results using buy-and-hold returns show that hedge returns for the media model are relatively larger, with a magnitude of 6.5% (t-statistic = 2.76). Similar results are presented in the last three columns using one-year-ahead abnormal returns, where the hedge returns are significant in the media model (4.1%, t-statistic = 2.37) but insignificant in the other two models.

To show, in terms of statistical significance, that the price deviation ratio from the media model has better return predictive ability than the other two models, we regress future returns on the decile ranking of price deviation ratios from each model. We control for various return predictors, including accruals, firm size, book-to-market ratio, institutional holdings, and buy-and-hold returns. Industry and year fixed effects are included. Standard errors are two-way clustered by firm and year. The results in Panel B show that the coefficients on *Price Deviation Ratio_Media Model* ^{dec} are significant in both columns while the coefficients on *Price Deviation Ratio_Simple Model* ^{dec} and *Price Deviation Ratio_AF Model* ^{dec} are insignificant. This evidence supports our conjecture that the media model outperforms the other two models.

Table 5 Predictive Ability of Price Deviation Ratios to Future Stock Returns

Panel A: Portfolio analysis

	Return$_{t+1}$			*AbRet$_{t+1}$*		
Rank based on *Price Deviation Ratio*	Simple Model	AF Model	Media Model	Simple Model	AF Model	Media Model
Low (underprice)	0.127	0.137	0.135	0.003	0.009	0.004
2	0.127	0.118	0.117	0.005	-0.002	-0.006
3	0.123	0.118	0.133	-0.000	0.001	0.007
4	0.131	0.102	0.117	0.011	-0.014	0.001
5	0.123	0.109	0.119	0.009	-0.004	0.005
6	0.128	0.106	0.118	0.019	-0.008	0.003
7	0.121	0.113	0.120	0.000	0.009	0.013
8	0.117	0.127	0.128	0.008	0.010	0.013
9	0.108	0.172	0.137	0.003	0.057	0.031
High (overprice)	0.087	0.092	0.070	-0.025	-0.022	-0.037
Hedge (Low-High)	0.040*	0.045*	0.065**	0.029	0.031	0.041**
	(1.90)	(1.87)	(2.76)	(1.55)	(1.61)	(2.37)

Panel B: Regression analysis

	Dep Var= *Return* $_{t+1}$	Dep Var = *AbRet* $_{t+1}$
Variable	(1)	(2)
Price Deviation Ratio_Simple Model^{dec}	-0.009	-0.004
	(-0.29)	(-0.19)
Price Deviation Ratio_AF Model^{dec}	0.042	0.025
	(1.39)	(0.96)
Price Deviation Ratio_Media Model^{dec}	-0.058**	-0.041*
	(-2.09)	(-1.82)
Accruals	-0.091	-0.064
	(-1.43)	(-0.93)
Size	0.005	0.003
	(1.35)	(1.10)
BTM	0.026	0.004
	(0.88)	(0.20)
Inst Holding	-0.136***	-0.139***
	(-6.85)	(-7.30)
Return	-0.025**	-0.025***
	(-2.33)	(-4.11)
Constant	-0.092**	0.027
	(-2.57)	(1.00)
Industry and Year Fixed Effect	Yes	Yes
Observations	20,534	20,544
Adjusted R-squared	0.1977	0.0144

This table examines the association between price deviation ratio and future abnormal returns. Panel A reports the results of the zero-investment portfolios. The sample consists of 21,338 firm-year observations. Portfolios are formed annually by assigning firms into deciles based on the *Price Deviation Ratio*, calculated by subtracting implied price model value from the market price, divided by market price. Portfolio returns are measured by $Return_{t+1}$ and $AbRet_{t+1}$. The hedge portfolio takes a long position in the lowest decile and a short position in the highest decile of price deviation ratio. T-statistics of hedge returns based on the time-series of annual portfolio abnormal stock returns are reported in parentheses. Panel B reports the results of regressing future returns on the price deviation ratio decile from three models. The sample consists of 20,534 firm-years in column (1) and 20,544 firm-years in column (2). *Price Deviation Ratio*^{dec} are the decile rankings of *Price Deviation Ratio* ranging from zero to one. Decile rankings are determined every year on the basis of the magnitude of *Price Deviation Ratio.* Industry and year fixed effects are included. T-statistics in brackets are based on two-way clustering by industry and year. See Appendix A for definitions of variables. *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

4.4 Impact of Media Characteristics

Next, we examine the mechanisms through which the media provides value-relevant information. We examine the effect of several key media characteristics on the performance of the media model in predicting abnormal earnings and stock price. We consider four media characteristics: deviation of news article sentiment from the contemporaneous firm press releases (*Media Difference*), sentiment dispersion of news articles (*Media Dispersion*), the number of news articles covering a firm (*Media Coverage*), and the frequency of news that is more relevant to firm performance (*Hard News Freq*).[7]

Dependent variables measure the superiority of the media model over the competing models in terms of accuracy in predicting abnormal earnings or price, including the difference in absolute FE and the difference in squared FE. A lower value of the dependent variables indicates the higher superiority of the media model over the competing model. We estimate the OLS regressions on the association between media characteristics and the media model superiority after controlling for firm fundamentals (*Size*, *BTM*) and some key aspects of information environment (*Inst Holding*, *Analyst Coverage*, *#Mgmt Guidance*). Industry and year fixed effects are included. The model is estimated using pooled data with standard errors clustered at industry and year level.

Table 6 reports the results on the superiority of the media model in predicting abnormal earnings. Columns (1) and (2) report the results of competing the media model with the simple model. Higher *Media Dispersion* and *Media Coverage* are associated with lower absolute FE difference and squared FE difference. This result suggests that more dispersed media sentiments and higher coverage of news articles contain more information about firm fundamentals that are incremental to earnings and book value. The coefficients on *Media Difference* and *Hard News Freq* are positive and significant at the 10% level in column (1) but become insignificant in column (2), which suggests some weak evidence on the role of the deviation of media sentiment from the press releases and the percentage of hard news in affecting the informativeness of media. Columns (3) and (4) compare the performance of the media model and the analyst forecast model. The negative and significant coefficient on *Media Dispersion* in column (4) shows limited evidence that the dispersion of media sentiment contributes to the media informativeness incremental to analyst forecasts.

We then extend the regression analysis to the price model. Table 7 presents the results of comparing the media model and the other models in terms of accuracy of predicting stock prices. Recall that lower values of dependent variables indicate higher superiority of the media model over the competing models. *Media Dispersion* and *Media Coverage* are generally negatively associated with FE difference and/or squared FE difference across four columns. This outcome suggests that dispersed media sentiments and higher media coverage positively contribute to the informativeness of media incremental to earnings, book value, and analyst forecast. In addition, we find that *Media Difference* has a negative impact on the informativeness of media, as evidenced by the positive and significant coefficient across four columns. There is also some evidence that *Hard News Freq* negatively contributes to the

[7] Following Wang *et al.* (2018), we utilise the news group categories in RavenPack and divide them into a hard news group and a soft news group. The hard news group is defined as being more relevant to firm fundamentals and thus consists of four news categories: revenues, earnings, analyst ratings, and credit ratings. All other news categories are included in the soft news group. The distribution of news categories in our sample is shown in Appendix B, and 14.51% of news is classified as hard news and the rest 85.47% are soft news. Among all business news articles, the top five news categories are “insider-trading” (63.34%), earnings (11.37%), labour issues (4.24%), product services (3.98%), and revenues (2.38%).

informativeness of the media incremental to analyst forecasts (0.075 in column (3), t-statistic = 2.08). This result suggests that soft news is likely to contain more incremental information about firm fundamentals than hard news.

Overall, the results in tables 6 and 7 suggest that both the information dissemination role (proxied by *Media Coverage*) and information creation role (proxied by *Media Dispersion*) of the media contain information that is incremental to earnings, book value, and analyst forecasts. The information dissemination role can highlight the importance of certain news related to a firm. The information creation role of the media can generate information that covers more topics and provides different opinions about a firm that provide information about a firm's fundamental value incremental to analyst forecasts. These results support our conjecture that the superiority of the media model stems from the incremental information contained in news articles relative to earnings, book values, and analyst forecasts.

Table 6 The Effect of Media Characteristics on the Performance of the Earnings Model

	Media Model vs Simple Model		Media Model vs AF Model	
Dep Var =	Absolute FE difference	Squared FE difference	Absolute FE difference	Squared FE difference
Variables	(1)	(2)	(3)	(4)
Media Difference	0.029*	0.007	-0.006	0.004
	(1.73)	(0.85)	(-0.28)	(0.42)
Media Dispersion	-0.059***	-0.018***	-0.006	-0.022**
	(-4.78)	(-4.07)	(-0.35)	(-2.21)
Media Coverage	-0.001***	-0.001**	-0.000	-0.000
	(-2.68)	(-2.21)	(-0.72)	(-0.97)
Hard News Freq	0.004*	0.001	0.000	0.001
	(1.82)	(0.85)	(0.15)	(0.48)
Inst Holding	-0.001	0.001***	0.003*	-0.002*
	(-0.72)	(2.59)	(1.75)	(-1.73)
Analyst Coverage	-0.002**	-0.001***	-0.001	-0.001**
	(-2.41)	(-3.03)	(-1.05)	(-2.22)
#Mgmt Guidance	0.000	-0.000	-0.001	0.000
	(0.52)	(-0.07)	(-1.51)	(0.16)
Size	0.002***	0.001***	0.000	0.001***
	(4.33)	(4.53)	(0.40)	(3.28)
BTM	-0.008***	-0.004***	-0.007***	-0.010***
	(-5.99)	(-5.96)	(-2.71)	(-4.28)
Constant	-0.007***	-0.003***	-0.001	0.000
	(-3.74)	(-2.95)	(-0.57)	(0.29)
Industry and Year Fixed Effect	Yes	Yes	Yes	Yes
Observations	22,648	22,648	22,648	22,648
Adj R-squared	0.0283	0.0289	0.0117	0.0317

This table presents the results for the effect of media characteristics on the superiority of the media model in predicting next year's abnormal earnings. The sample consists of 22,648 firm-years from 2000 to 2017. Media characteristics include the sentiment difference between news articles and press releases (*Media Difference*), the sentiment dispersion of news articles (*Media Dispersion*), the number of news articles (*Media Coverage*), and the frequency of hard news (*Hard News Freq*). Absolute FE and squared FE are two metrics that measure the accuracy of the predictions of next period abnormal earnings generated by each of the earnings valuation models. Absolute FE difference (Squared FE difference) is calculated as absolute FE (squared FE) from the media model minus absolute FE (squared FE) from the simple model in the first two columns and from the AF model in the last two columns. Industry and year fixed effects are included. T-statistics in brackets are based on two-way clustering by industry and year. See Appendix A for definitions of the variables. *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

Table 7 The Effect of News Characteristics on the Performance of the Price Model

	Media Model vs Simple Model		Media Model vs AF Model	
Dep Var =	Absolute FE difference	Squared FE difference	Absolute FE difference	Squared FE difference
Variables	(1)	(3)	(2)	(4)
Media Difference	0.487***	0.678***	1.014*	7.075*
	(4.49)	(3.12)	(1.95)	(1.95)
Media Dispersion	-0.642**	-0.636	-2.048***	-12.025***
	(-2.48)	(-0.93)	(-4.19)	(-3.77)
Media Coverage	-0.010**	-0.018***	-0.018	-0.132*
	(-2.55)	(-2.74)	(-1.31)	(-1.70)
Hard News Freq	0.009	-0.010	0.075**	0.362
	(0.39)	(-0.17)	(2.08)	(1.30)
Inst Holding	-0.090***	-0.073**	0.102***	-0.106
	(-4.30)	(-2.40)	(2.93)	(-0.50)
Analyst Coverage	0.006	0.017	-0.053***	-0.317***
	(0.49)	(0.52)	(-3.69)	(-3.30)
#Mgmt Guidance	-0.007	-0.018	0.005	0.005
	(-1.13)	(-1.28)	(0.47)	(0.09)
Size	0.007	0.002	0.052***	0.313***
	(0.61)	(0.05)	(4.28)	(4.70)
BTM	0.076***	0.109**	-0.489***	-1.815***
	(3.04)	(2.25)	(-5.06)	(-3.17)
Constant	-0.052	-0.026	-0.076	-0.607**
	(-0.73)	(-0.14)	(-1.25)	(-1.98)
Industry and Year Fixed Effect	Yes	Yes	Yes	Yes
Observations	22,648	22,648	22,648	22,648
Adj R-squared	0.0442	0.0348	0.1565	0.0780

This table presents the results for the effect of media characteristics on the superiority of the media model in predicting next year's stock price. The sample consists of 22,648 firm-years from 2000 to 2017. Media characteristics include the sentiment difference between news articles and press releases (*Media Difference*), the sentiment dispersion of news articles (*Media Dispersion*), the number of news articles (*Media Coverage*), and the frequency of hard news (*Hard News Freq*). Absolute FE and squared FE are two metrics that measures the accuracy of the predictions of next period stock prices generated by each of the earnings valuation models. Absolute FE difference (Squared FE difference) is calculated as absolute FE (squared FE) from the media model minus absolute FE (squared FE) from the simple model in the first two columns and from the AF model in the last two columns. Industry and year fixed effects are included. T-statistics in brackets are based on two-way clustering by industry and year. See Appendix A for definitions of the variables. *, **, and *** indicate significance levels at 10%, 5%, and 1%, respectively.

V. Conclusion

In this study, we examine whether media sentiment contains information about a firm's fundamental value incremental to earnings, book value, and analyst forecasts. Using Ohlson's

(1995) RIM, we show that incorporating media content into the model generally improves its ability to predict future residual income, explain current stock prices, and predict future stock prices. We also show that our results are stronger when there are more news articles covering the firm and when the news sentiments are more dispersed.

This study adds to the fundamental valuation literature by exploring a new method to measure the information dynamics of the other information in Ohlson's (1995) model. Most empirical studies that implement Ohlson's (1995) model face challenges in measuring the other information, and scholars mainly rely on using analyst forecasts to measure it (e.g. Frankel and Lee, 1998; Dechow *et al.*, 1999; Bryan and Tiras, 2007). Our study uses media data to improve the measurement of the other information.

This study also contributes to the literature on the media and the capital market by examining the media as a source of valuation. Diverging from prior studies that examine the association between the media and investors' responses to accounting information or information asymmetry (Bushee *et al.*, 2010; Drake *et al.*, 2014), we use the RIM to examine whether media information improves the reliability of intrinsic value estimates derived from Ohlson's model. In addition, while prior research focuses on media coverage (Bushee *et al.*, 2010; Drake *et al.*, 2014), we show that media content (measured by the sentiment of news articles), another key feature of the media, contains information incremental to accounting data and analyst forecasts. We also show that this effect is more pronounced when media coverage is higher and when media sentiment is more dispersed.

References

Barber, B. M. and Odean, T. (2008), 'All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors', *Review of Financial Studies* 21 (2): 785–818.

Baron, D. P. (2006), 'Persistent Media Bias', *Journal of Public Economics* 90 (1-2): 1–36.

Beaver, W. H. (1999), 'Comments on "An empirical assessment of the residual income valuation model"', *Journal of Accounting and Economics* 26 (1-3): 35–42.

Bradshaw, M., Lock, B., Wang, X., and Zhou, D. (2021), 'Soft Information in the Financial Press and Analyst Revisions', *The Accounting Review* 96 (5): 107–132.

Brief, R. P. (2007), 'Accounting valuation models: A short primer', *Abacus* 43 (4): 429–437.

Brown, L. D. (1991), 'Forecast Selection When All Forecasts Are Not Equally Recent', *International Journal of Forecasting* 7 (3): 349–356.

Bryan, D. M. and Tiras, S. L. (2007), 'The Influence of Forecast Dispersion on the Incremental Explanatory Power of Earnings, Book Value, and Analyst Forecasts on Market Prices', *The Accounting Review* 82 (3): 651–677.

Bushee, B. J., Core, J. E., Guay, W., and Hamm, S. J. W. (2010), 'The Role of the Business Press as an Information Intermediary', *Journal of Accounting Research* 48 (1): 1–19.

Bushman, R. M., Williams, C. D., and Wittenberg-Moerman, R. (2017), 'The informational role of the media in private lending', *Journal of Accounting Research* 55 (1): 115–152.

Cao, Y., Keskek, S., Myers, L. A., and Tsang, A. (2020), 'The Effect of Media Competition on Analyst Forecast Properties: Cross-Country Evidence', *Journal of Accounting, Auditing, and Finance*, forthcoming.

Chen, Y., Cheng, C. S. A., Li, S., and Zhao, J. (2021), 'The monitoring role of the media: Evidence from earnings management', *Journal of Business Finance & Accounting* 48 (3-4): 533–563.

Clarkson, P. M., Li, Y., and Richardson, G. D. (2004), 'The Market Valuation of Environmental Capital Expenditures by Pulp and Paper Companies', *The Accounting Review* 79 (2): 329–353.

Core, J. E., Guay, W., and Larcker, D. F. (2008), 'The Power of the Pen and Executive Compensation', *Journal of Financial Economics* 88 (1): 1–25.

Damodaran, A. (2020), 'Equity Risk Premiums: Determinants, Estimation and Implications—The 2020 Edition', available at https://ssrn.com/abstract=3550293.

DeAngelo, H., DeAngelo, L., and Gilson, S. C. (1994), 'The Collapse of First Executive Corporation Junk Bonds, Adverse Publicity, and the "Run on the Bank" Phenomenon', *Journal of Financial Economics* 36 (3): 287–336.

DeAngelo, H., DeAngelo, L., and Gilson, S. C. (1996), 'Perceptions and the Politics of Finance: Junk Bonds and the Regulatory Seizure of First Capital Life', *Journal of Financial Economics* 41 (3): 475–511.

Dechow, P. M., Hutton, A. P., and Sloan, R. G. (1999), 'An Empirical Assessment of the Residual Income Valuation Model', *Journal of Accounting and Economics* 26 (1-3): 1–34.

Dong, M., Hirshleifer, D., Richardson, S., and Teoh, S. H. (2006), 'Does Investor Misvaluation Drive the Takeover Market?', *The Journal of Finance* 61 (2): 725–762.

Drake, M. S., Guest, N. M., and Twedt, B. J. (2014), 'The Media and Mispricing: The Role of the Business Press in the Pricing of Accounting Information', *The Accounting Review* 89 (5): 1673–1701.

Engelberg, J. E. and Parsons, C. A. (2011), 'The Causal Impact of Media in Financial Markets', *The Journal of Finance* 66 (1): 67–97.

Fang, L. and Peress, J. (2009), 'Media Coverage and the Cross-Section of Stock Returns', *The*

Journal of Finance 64 (5): 2023–2052.

Foster, G. (1979), 'Briloff and the capital market', *Journal of Accounting Research* 17 (1): 262–274.

Frankel, R. and Lee, C. M. C. (1998), 'Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns', *Journal of Accounting and Economics* 25 (3): 283–319.

Gentzkow, M. and Shapiro, J. M. (2010), 'What Drives Media Slant? Evidence From U.S. Daily Newspapers', *Econometrica* 78 (1): 35–71.

Gode, D. and Ohlson, J. (2004), 'Accounting-Based Valuation with Changing Interest Rates', *Review of Accounting Studies* 9 (4): 419–441.

Gu, C. and Kurov, A. (2020), 'Informational role of social media: Evidence from Twitter sentiment', *Journal of Banking & Finance* 121: Article 105969.

Jensen, M. C. (1979), 'Toward a Theory of the Press', in Brunner, K. (ed.), *Economics Social Institutions* (Rochester Studies in Economics and Policy Issues, Vol. 1), Dordrecht, Netherlands: Springer, pp. 267–287.

Jorgensen, B. N., Lee, Y. G., and Yoo, Y. K. (2011), 'The valuation accuracy of equity value estimates inferred from conventional empirical implementations of the abnormal earnings growth model: US evidence', *Journal of Business Finance and Accounting* 38 (3-4): 446–471.

Klibanoff, P., Lamont, O., and Wizman, T. A. (1998), 'Investor reaction to salient news in closed-end country funds', *The Journal of Finance* 53 (2): 673–699.

Kothari, S. P., Li, X., and Short, J. E. (2009), 'The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis', *The Accounting Review* 84 (5):1639–1670.

Kross, W., Ro, B., and Schroeder, D. (1990), 'Earnings Expectations: The Analysts' Information Advantage', *The Accounting Review* 65 (2): 461–477.

Lee, C. M. C., Myers, J., and Swaminathan, B. (1999), 'What Is the Intrinsic Value of the Dow?', *The Journal of Finance* 54 (5): 1693–1741.

Lo, K. and Lys, T. (2000), 'The Ohlson Model: Contribution to Valuation Theory, Limitations, and Empirical Applications', *Journal of Accounting, Auditing and Finance* 15 (3): 337–367.

Lundholm, R. and O'Keefe, T. (2001), 'Reconciling Value Estimates from the Discounted Cash Flow Model and the Residual Income Model', *Contemporary Accounting Research* 18 (2): 311–335.

McNichols, M. and O'Brien, P. C. (1997), 'Self-Selection and Analyst Coverage', *Journal of Accounting Research* 35 (Supplement): 167–199.

Mullainathan, S. and Shleifer, A. (2005), 'The Market for News', *The American Economic Review* 95 (4): 1031–1053.

Ohlson, J. A. (1995), 'Earnings, Book Values, and Dividends in Equity Valuation', *Contemporary Accounting Research* 11 (2): 661–687.

Ohlson, J. A. (2001), 'Earnings, book values, and dividends in equity valuation: An empirical perspective', *Contemporary Accounting Research* 18 (1): 107–120.

Ohlson, J. A. and Juettner-Nauroth, B. E. (2005), 'Expected EPS and EPS Growth as Determinants of Value', *Review of Accounting Studies* 10 (2-3): 349–365.

Penman, S. H. (2005), 'Discussion of "On Accounting-Based Valuation Formulae" and "Expected EPS and EPS Growth as Determinants of Value"', *Review of Accounting Studies* 10 (2-3): 367–378.

Penman, S. H. (2015), 'Valuation models: An issue of accounting theory', in *The Routledge Companion to Financial Accounting Theory*, London: Routledge, pp. 236–253.

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.

Rhodes-Kropf, M., Robinson, D. T., and Viswanathan, S. (2005), 'Valuation Waves and Merger Activity: The Empirical Evidence', *Journal of Financial Economics* 77 (3): 561–603.

Sant, R. and Zaman, M. A. (1996), 'Market Reaction to Business Week "Inside Wall Street" Column: A Self-Fulfilling Prophecy', *Journal of Banking & Finance* 20 (4): 617–643.

Soltes, E. (2010), 'Disseminating Firm Disclosures', Working Paper, Harvard University.

Tetlock, P. C. (2007), 'Giving Content to Investor Sentiment: The Role of Media in the Stock Market', *The Journal of Finance* 62 (3): 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008), 'More than Words: Quantifying Language to Measure Firms' Fundamentals', *The Journal of Finance* 63 (3): 1437–1467.

von Beschwitz, B., Keim, D. B., and Massa, M. (2020), 'First to "Read" the News: News Analytics and Algorithmic Trading', *The Review of Asset Pricing Studies* 10 (1): 122–178.

Wang, Y., Zhang, B., and Zhu, X. (2018), 'The Momentum of News', available at https://ssrn.com/abstract=3267337.

Zhao, J. (2017), 'Corporate Information Environment Dynamics and Equity Valuation—An Analysis of Mandatory Disclosures, Voluntary Disclosures, and Information Intermediaries', Working Paper, The Hong Kong Polytechnic University.

Appendix A: Variable Definitions

Variable	Definition
Valuation Model Variables	
P_t	Stock price measured at the end of fiscal year *t*
b_t	The book value of equity at the end of year *t*
x_t^a	Abnormal earnings, measured as $x_t - rb_{t-1}$, with *r* being the discount rate (assumed to be 8%)
ω^u	Unconditional ω is the first-order autoregression coefficient for abnormal earnings: $x_{t+1}^a = \alpha + \omega^u x_t^a + \varepsilon_{1,t+1.}$
ω^c	Conditional ω is the predicted value of ω, measured as $\omega_1 + \omega_2 q1_t + \omega_3 q2_t + \omega_4 q3_t + \omega_5 q4_t + \omega_6 q5_t$, where ω_1 to ω_5 are estimated from Eq (4). q1 is the absolute value of abnormal earnings for year *t* divided by book value of equity at the beginning of year *t*; q2 is the absolute value of special items divided by book value of equity at the beginning of year *t*; q3 is the absolute value of accruals divided by book value of equity at the beginning of year *t*; q4 is dividends paid during year *t* divided by earnings before extraordinary items for year *t*; and q5 is the first-order autoregressive coefficient from an abnormal earnings autoregression for all firms in the same two digit SIC code. The autoregression is conducted in the same two digit SIC code from 1960 to year *t*.
v_t	Other information, measured as $f_t^a - \omega^u x_t^a$
γ^u	Unconditional γ is the first-order autoregression coefficient for v_t: $v_{t+1} = \alpha + \gamma^u v_t + \varepsilon_{2,t+1}$
γ^c	Conditional γ is measured as $\gamma_1 + \gamma_2 m_t$, where γ_1 and γ_2 are estimated from Eq. (6).
f_t	The IBES consensus forecast of earnings for year t+1 measured in the first month following the announcement of earnings for year *t*
f_t^a	$f_t - r * b_t$, where *r* denotes the discount rate (assumed to be 8%)
m_t	Media content, measured as the average (CSS-50)/100 of all news articles for a firm-year. Firm-initiated press releases and news flashes are excluded from this estimation. Composite Sentiment Score (CSS) is a story-level sentiment analytic that represents news sentiment by combining various sentiment analysis techniques. See Bushman *et al.* (2017) appendix for more details regarding how RavenPack constructs the CSS.
Media Characteristics	
Media Difference	The absolute difference of the average sentiment score between press releases and media news for a firm-year
Media Dispersion	Dispersion in media sentiment, measured as the standard deviation of news article sentiments for a firm-year

Hard News Freq	Frequency of hard news, measured as the number of hard news articles divided by the number of total news articles. Hard news is defined as news more relevant to firm fundamentals. News categories of hard and soft news are shown in Appendix B.
Media Coverage	The aggregated number of business-press initiated news articles about a firm in year *t*, divided by 100
Other Variables	
$AbRet_{t+1}$	Buy-and-hold abnormal returns in year t+1, calculated over 12 months starting 4 months after the end of fiscal year t; size-adjusted return subtracts from each stock return the return on a portfolio of firms matched on size; DGTW return subtracts from each stock return the return on a portfolio of firms matched on market equity, market-book, and prior one-year return quintiles.
Price Deviation Ratio	Price forecast error ratio, measured as realised price minus implied price from price model, dividend by realised price
Inst Holding	Percentage of outstanding shares held by institutional owners
Return	Buy-and-hold returns in year *t*
BTM	Ratio of book value of common equity to market value
Size	The natural logarithm of market value at the end of year *t*
Analyst Coverage	Number of analysts whose forecasts are included in the most recent consensus before a firm's annual earnings announcement; if a firm-year is covered by Compustat but not by IBES, we code it as zero.
#Mgmt Guidance	Frequency of management guidance (quarterly and annual) issued in year *t*. Missing values are coded as 0.

Appendix B: List of News by Categories

All the business news articles in our study can be grouped into 22 news categories as follows. To differentiate news articles with respect to the amount of other information they contain, we follow Wang *et al.* (2018) to decompose news into hard news and soft news. The hard news group is defined as more relevant to firm fundamentals and thus consists of four news categories: revenues, earnings, analyst ratings and credit ratings. All other news categories are included in the soft news group.

Categories	News Group	Frequency (%)
Hard News	earnings	11.37
	revenues	2.38
	analyst ratings	0.51
	credit ratings	0.25
	Subtotal	14.51
Soft News	insider trading	63.34
	labour issues	4.24
	product services	3.98
	stock prices	3.67
	acquisitions & mergers	3.53
	equity actions	2.21
	legal	1.23
	assets	0.72
	credit	0.71
	investor relations	0.55
	dividends	0.37
	partnerships	0.37
	regulatory	0.22
	price targets	0.14
	marketing	0.05
	bankruptcy	0.04
	industrial accidents	0.03
	exploration	0.02
	security	0.02
	government	0.01
	indexes	0.01
	war conflict	0.01
	balance of payments	0
	civil unrest	0
	corporate responsibility	0
	crime	0
	order imbalances	0
	pollution	0
	public opinion	0
	taxes	0
	technical analysis	0
	transportation	0
	Subtotal	85.47