Good News is No News: Asymmetric Inattention and the Neglected Firm Effect

Charles Gaa[*]

August 1, 2008

Abstract

This paper identifies asymmetric investor attention as a potential explanation for the neglected firm premium in the cross-section of stock returns. Utilizing ex ante predicted probability of media coverage (PMC) with respect to earnings news as a measure of investor attention, I show that a portfolio strategy based on long positions in low-attention stocks yields excess returns of approximately 70 bps per month between 1984 and 2005. The finding of high risk-adjusted returns for these "neglected" stocks (typically small firms with low analyst coverage, low institutional ownership, and little trading activity) appears to be highly consistent with, e.g., Merton's (1987) investor recognition hypothesis, or an information risk setting (Easley et al. (2002)). However, in examining the event-specific determinants of media coverage, I find evidence of a significant "negativity bias" in attention: holding other factors constant, bad news is more likely to attract coverage than is good news regarding an otherwise-identical firm (i.e., "if it bleeds, it leads"). Given recent evidence in the literature regarding stock-price underreaction to low-attention events, this suggests that findings of significant premia for "neglected" stocks may be explained, at least in part, by an asymmetric underreaction to positive news events from these firms. Consistent with this hypothesis, I find that the excess returns to low-PMC portfolios are attributable to drift in the stock prices of low-attention "good news" firms, while low-attention "bad news" firms appear to be efficiently priced.

JEL classifications: G14; L82; D83
Keywords: media coverage; stock returns; earnings announcements; investor attention; neglected firms

---

[*] PhD candidate, Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2. Email: charles.gaa@sauder.ubc.ca. [Previous working title: "Media Coverage, Investor Inattention, and the Market's Reaction to News".] I would like to thank my committee: Kai Li (Chair), Werner Antweiler, Adlai Fisher, and Marcin Kacperczyk for their ongoing support and valuable feedback. I am also grateful to Murray Carlson, Jason Chen, Alex Edmans, Maurice Levi, Hernan Ortiz-Molina, and UBC faculty members and seminar participants for helpful comments and suggestions. All errors are mine.

1 Introduction

In attempting to explain persistent findings of high risk-adjusted returns for "neglected", low-recognition stocks,[1] researchers have typically focused on the identification and estimation of additional sources of risks and/or frictions that may not be fully priced in the traditional asset pricing models. In particular, one influential line of research has pointed to problems in the information environments faced by these firms (e.g., Merton (1987), Easley et al. (2002)). In this context, the observation of excess returns for neglected/low-information stocks can be explained as compensation for increased costs of information acquisition, exacerbated parameter uncertainty, or heightened asymmetric information, etc. Recently, researchers have also begun to focus on a related set of information characteristics (e.g., analyst coverage, institutional ownership, trading volumes, media coverage, etc.) as proxies for investor attention in attempting to explain patterns of apparent underreaction (and overreaction) in stock returns (Hou et al. (2006), Barber and Odean (2008), Chan (2003), Brennan et al. (1993), Hong et al. (2000)). This paper creates a link between these two strands of the literature by identifying underreaction to positive news related to asymmetric investor attention (i.e., "negativity bias") as a potential explanation for an apparent neglected firm premium in stock returns.

Constrained attention effectively prevents investors from acquiring and processing all of the potentially-relevant information that might be available at any given point in time. But how do agents decide which information items are worthy of attention, and which others will, of necessity, be ignored? While theoretical models predict that limited investor attention can lead to predictability in asset prices (see Peng and Xiong (2006) or Huang and Liu (2007)), attention allocations are extremely difficult to observe in practice. As a result, our understanding of this potentially important area of economic decision-making has continued to represent something of a "black box" for researchers.

This paper examines financial news reporters' coverage decisions in order to identify the event-specific (as well as the firm-specific and market-wide) factors that predict whether a particular earnings news event will be communicated to investors via this highly important information channel. Utilizing ex ante predicted probability of media coverage (PMC) as a new measure of investors' attention allocations in this context, I find that portfolio strategies with long positions in low-PMC stocks generate excess returns of approximately 70 bps per month after controlling for the standard risk factors identified in the literature. Insofar as this finding is

[1] See, e.g., evidence regarding firm size and analyst coverage (e.g., Arbel and Strebel (1982)), cross-listing (Foerster and Karolyi (1999)), "delayed" firms (Hou and Moskowitz (2005)), and financial news media coverage (Fang and Peress (2007)).

consistent with the existing evidence regarding premia for "neglected" and "delayed" firms, it should not be surprising. Upon closer examination of the event-specific determinants of coverage, however, an alternative explanation emerges. In particular, I find evidence of a significant "negativity bias" in media attention: Bad news is more likely to result in coverage than is good news regarding an otherwise-identical firm. Given recent empirical evidence that market prices systematically underreact to low-attention events, and to news from low-attention firms in general, asymmetric underreaction to positive news emerges as a potential alternative to the standard information friction- and risk-based explanations for the neglected firm effect. Consistent with the asymmetric underreaction hypothesis, I find that the observed excess returns to low-PMC portfolios are attributable to high returns for low-attention "good news" firms, while low-attention "bad news" firms appear to be efficiently priced.

A growing body of research indicates that the mainstream financial news media, in particular, represents an important source of information in financial markets.[2] One explanation simply refers to the relative costs of information acquisition and processing: information in the news media is cheap and is typically presented in way that it is quickly and easily understood by non-specialists. Another, potentially more interesting, interpretation for the significance of media information points to the fact that news media outlets actively compete with each other in attempting to anticipate reader interest – in other words, news reporters can be seen as effectively possessing a "theory of mind" regarding their customers' underlying preferences for new information. Given investors' cognitive constraints, then, it seems natural that the financial news media would play a crucial role as an information intermediary, providing relatively cheap and easy access to a sub-set of news items that the typical consumer will, on average, find most useful and interesting. I argue that financial news media coverage decisions provide a crucial window for our understanding of the determinants of investor attention. Put simply, events and firms whose characteristics predict a greater likelihood of receiving financial news media coverage are expected to attract higher levels of attention from investors.

To the present day, researchers continue to find evidence of high unexplained returns related to various proxies for the quality of a firm's information environment: the neglected firm effect. In this context, the concept of neglect is potentially quite broad. Arbel and Strebel (1982) were among the first to identify such a premium by looking at the relationship between stock returns and the number of securities analysts following a firm. More recently, Hou and

[2] See, e.g., Huberman and Regev (2001), Busse and Green, (2002), Chan (2003), Dyck and Zingales (2003), Barber and Odean (2008), Bhattacharya et al. (2006), Tetlock (2007), Tetlock et al. (2008), Antweiler and Frank (2006), and Fang and Peress (2007).

Moskowitz (2005) show that firms whose stock prices exhibit significant "delay" with respect to their adjustment to common information shocks subsequently experience high returns that cannot be explained by the standard set of risk factors identified in the literature. Looking at the potential role of media coverage, in particular, Fang and Peress (2007) find that firms with no recent media coverage significantly outperform those who have experienced relatively high levels of coverage. The authors of both of these papers suggest that the identified return premia are consistent with the effects of frictions in the information environments faced by these firms. More broadly, depending upon one's definition of neglect, investor recognition (Merton (1987), Shapiro (2002), Basak and Cuoco (1998)), information risk (Easley et al. (2002)), and illiquidity (Amihud and Mendelson (1986), Pastor and Stambaugh (2003)) are all potentially consistent with the observation of a neglected firm premium in the cross-section of stock returns.

While low attention (or neglect) has traditionally been viewed primarily as a potential source of additional friction and/or risk for which investors must be compensated, researchers have also begun to examine investor attention as a potential explanation for observed variation in the speed with which different firms' stock prices react to news events. For example, Huang and Liu (2007) show that inattention to public news is rational when information acquisition is costly, potentially leading to over- or under-investment in portfolio selection when signals are noisy. If limits to arbitrage are sufficient to prevent prices from quickly adjusting to eliminate mispricing, then there is scope such phenomena to affect market prices, at least in the short term (Delong et al. (1990)). Consistent with delayed price reactions to news from neglected firms, Brennan et al. (1993) and Hong, Lim, and Stein (2000) (HLS) provide evidence that the stock prices of firms with low analyst coverage appear to respond more sluggishly to information shocks.

In this context, earnings announcements provide a particularly good testing ground for a study investigating the relationship between investor attention and potential underreaction in the market's response to an underlying news event. Going back as far as Ball and Brown (1968), post-earnings-announcement drift (PEAD) – the tendency for stock returns to exhibit continuation following an earnings surprise – has represented a significant puzzle for researchers. Bernard and Thomas (1990) ask whether PEAD represents a risk premium or a delayed response (i.e., underreaction) to earnings news; they find that their results are more consistent with the latter interpretation. If PEAD can be explained as an underreaction to earnings news, it is natural to ask whether, as one might expect, the phenomenon is stronger when investor attention is expected to be low. Indeed, there have been several studies linking the extent and/or speed with which prices react to earnings news to a number of potential proxies for attention, such as: firm size (Bamber (1987), Christensen et al. (2004)), trading volumes and overall market conditions (Hou et al.

(2006)), analyst coverage (Christensen et al. (2004)), the presence of competing news events (DellaVigna and Pollet (2008), Hirshleifer et al. (2006)), and the ease with which earnings information may be processed (Engelberg (2007)).

In particular, there is recent research regarding the impact of media coverage on the market's reaction to earnings news. For example, Dyck and Zingales (2003) show that prices appear to react more strongly to the earnings numbers emphasized by reporters in their stories. Moreover, a large proportion of overall firm-specific media coverage seems to be linked to earnings news flows (Tetlock et al. (2007)), and they are often seen as the single most important piece of information regarding a firm's performance and future prospects. Tetlock (2008) shows that coverage surrounding earnings announcements contains significant incremental information regarding subsequent stock returns, while Tetlock et al. (2008) demonstrate the link to firms' fundamentals. Furthermore, Engelberg (2007) finds that the "soft", qualitative earnings information in news stories contains significant information regarding future returns, consistent with delayed reaction due to increased processing costs.

Given the aforementioned research, what is the contribution of this paper? While prior studies focus on the information content of earnings news coverage and its potential impact on stock returns, I examine financial reporters' story selection decision-making as an indicator of investor attention: Which earnings news events are most likely to be covered? In estimating these coverage decisions, I seek to identify the firm, event, and market-level characteristics that are associated with high levels of investor attention, and, conversely, "neglect". I demonstrate that negative news is more likely to be covered than positive news, ceteris paribus, implying that attention (and, therefore, the market's reaction) is potentially endogenous with respect to the information content of the underlying event itself. Finally, I study the potential impact of this negativity bias in the context of PEAD, presenting evidence that an apparent neglected firm premium in this context is actually more consistent with asymmetric underreaction to positive news from low-attention firms.

This paper's finding that attention varies with the information content of the underlying event is broadly consistent with evidence in the literature regarding asymmetric responses to positive and negative news (e.g., McQueen et al. (1996), Kothari et al. (2005), Veronesi (1999), Conrad et al. (2002), and Skinner and Sloan (2002)). In particular, researchers have documented an apparent negativity bias with respect to the media's coverage of macroeconomics news. Harrington (1989), for example, finds that the U.S. media generally pays greater attention to bad economic news, particularly in non-election years. More recently, Soroka (2006) shows that U.K news media and public opinion are more responsive to negative macroeconomic news releases than positive ones. For the first time, I present similar findings with respect to the financial news media's coverage of corporate earnings announcements.

But why would investors choose to allocate more attention to negative events? While this question is somewhat beyond the scope of the current paper, some discussion is warranted. Negativity bias, as a general phenomenon, is seen as consistent with prospect theory and loss aversion (Kahneman and Tversky (1979)). In particular, loss-averse agents are willing to expend more resources in avoiding a loss than in pursuing an equivalent-sized gain. If attention is a scarce resource, then it makes sense for loss-averse investors to expend relatively more attention in monitoring negative news, since negative events may be more likely to contain information regarding potential losses that may affect them. However, there are also other, more "rational" explanations that point to the underlying nature of the problem faced by investors in interpreting financial and accounting data in this context. In particular, if investors perceive that managers are reluctant to communicate bad news (due to career concerns, or in order to maximize the near-term value of their equity holdings in the firm, for example), it follows that markets would react more strongly to an unambiguously negative signal, inferring that it is more likely to be truthful than a positive one, on average. Also, if managers tend to engage in "big bath" accounting, then we might expect relatively infrequently-observed negative news announcements to elicit asymmetrically large responses from investors. For example, Kothari et al. (2005) argue that managers tend to withhold bad news; consistent with this, they find that price reactions to bad news disclosures are significantly larger than those precipitated by positive ones. Similarly, Skinner and Sloan (2002) argue that asymmetrically large reactions to negative earnings announcements for growth stocks are linked to periodic release of disappointing announcements in the face of baseline expectations that tend to be overoptimistic with respect to these firms.

The rest of the paper is organized as follows. Section 2 motivates and develops the methodological strategy employed in this study. Section 3 describes the data set. Section 4 investigates the firm- and event-specific determinants of media coverage. Section 5 describes the construction of my proposed measures of investor attention based on the predictable component of media coverage decisions. Section 6 examines the potential implications of biased inattention for the cross-section of stock returns and the market's reaction to news. Section 7 describes a series of robustness tests on the main predictability findings. Section 8 concludes.

2 Methodology and Hypothesis Development

In this section, I develop and motivate the methodology of the paper. First, I apply tools from computational linguistics to identify a large data set of news articles that specifically discuss the results of corporate earnings announcements, and then subsequently to estimate the tone of coverage, if any. Second, linking observed instances of both positive and negative media coverage to the underlying corporate information releases that precipitate them, I explain media coverage decisions in reference to event- and firm-specific story elements. Third, I employ probability of media coverage (PMC) as a proxy for investors' realized attention allocations, or, alternatively, as a measure of their underlying information consumption preferences, and investigate the factors contributing to an apparent neglected firm premium in this context.

Determining the subject and tone of media coverage

A primary goal of the paper is to identify the precise story elements that predict positive and negative media coverage with respect to specific corporate earnings announcements. However, given the enormous quantity of potentially-relevant coverage available in news media databases, the task of locating articles of this very specific type, and of then classifying them according to semantic content, can present a technical challenge. Fortunately, this type of problem is not a new one in the field of computational linguistics, and a suite of techniques and tools have been developed (see, e.g., Manning and Schuetze (1999)).

In creating the media data set, I follow Antweiler and Frank (2004) in applying a Naïve-Bayesian "bag-of-words" approach using the Rainbow Toolkit (McCallum (1996)). A brief description follows (for a more detailed explanation, see Appendix A). In the first stage, a relatively small "training set" of 500 articles is randomly selected from the larger set and classified into semantic categories by hand. Next, the Rainbow text classification program builds an empirical model of class membership based on observed word frequencies for articles from the training set. ("Bag-of-words" refers to the fact that word order is not considered, only frequency of occurrence.) Each word that appears in a training set article is assigned an odds ratio based on its ability to predict membership in each semantic category. Finally, in order to classify text documents outside of the original training set, the program calculates sums of the odds ratios corresponding to the words found in each document, assigning documents to the category classification that maximizes the sum.

Which factors predict media coverage?

In this paper, I investigate the firm-specific and announcement-specific factors that predict media coverage, as well as the tone of coverage, if any. An essential requirement, therefore, is a large sample of information events with the potential to attract coverage. There must be variation in terms of coverage, and I need to observe the characteristics of events that do not receive coverage as well as those that do (i.e., related media coverage cannot be the only source of information regarding the event in question). Ideally, event characteristics should also be quantifiable and comparable across firms and over time. Since we are interested here in potential asymmetries in the attention paid to events of different kinds, the event in question should be reasonably classifiable as relatively "good" or "bad" news based on some metric that is, in particular, independent of the market's price reaction and the observed tone of coverage, if any.

In short, corporate earnings announcements provide a nearly-ideal setting for a study of this kind. Publicly-listed firms are typically required to produce quantitative accounting information on a quarterly basis, no matter whether that information might be viewed as good or bad, interesting or boring – this essentially eliminates the potential problem of endogenous information production and dissemination by firms. Earnings announcements also represent a very large set of information events, each one of which has the potential to attract media coverage ex ante (although, ex post, we know that most of them do not). Crucially, we can also interpret a particular firm's earnings release as being relatively positive or negative relative to other firms' announcements by, e.g., comparing its stated performance to analysts' pre-announcement forecasts, or simply by classifying the release as a "profit" or a "loss" in absolute terms. Therefore, in addition to non-event-specific predictors of firm coverage (e.g., as identified by Fang and Peress (2007) and Engelberg (2007)), I consider a set of factors related to the content of the information event itself. Furthermore, it is important to ensure that all of the proposed predictors of coverage represent information that was potentially available before the coverage decision was made; in particular, using contemporaneously-observed event-window CARs (rather than the earnings surprise relative to expectations) to predict coverage raises potential worries with respect to reverse-causality.

Given this data set of quarterly earnings announcements, which includes information on market conditions, the information content of the earnings releases, as well as characteristics of the firms making the announcements, I attempt to link each of the events to an identified earnings announcement-related news article in the Wall Street Journal (WSJ) regarding that firm. Approximately 13% of earnings announcements in the sample can be linked to an observation of related, contemporaneous media coverage.[3] I first examine the impact of event characteristics on the absolute probability of coverage by performing logit regressions with a dummy variable

[3] The relatively selective nature of coverage here (elsewhere, e.g., Engelberg (2007) finds that approximately one-half of the announcements in his sample are covered by Dow Jones News Service) may reflect the fact that the WSJ stories used in this study are all "stand-alone" articles, written by reporters themselves, with the subjects presumably chosen with a closer eye to potential reader interest, while newswire stories are often summaries (or even reproductions) of company-produced press releases. Given the potential selection issues related to firm-originated news, this may also help to explain the relatively strong evidence of asymmetric coverage (i.e., negativity bias) presented in this paper.

representing WSJ coverage on the left-hand-side and potential determinants of coverage on the right-hand-side. Subsequently, I investigate potential asymmetry regarding the impacts of story elements on the tone of coverage by differentiating observations of positive and negative coverage in a multinomial logit setting.

Media coverage as a proxy for investor attention and neglect

Having identified a set of factors that are significant in explaining variation in the probability and tenor of media coverage, I use predicted values from the multinomial logit estimation as a proxy for relative attention and neglect in the overall information environment. I then proceed to investigate the neglected firm effect in this context. While somewhat novel in its construction, I argue that this interpretation of neglect parallels and complements prior definitions based on characteristics such as trading activity, size, analyst coverage, etc.

Under what kinds of conditions might estimated media coverage represent an appropriate proxy for investor attention in a broader sense? First, if the Wall Street Journal (WSJ) is, in and of itself, a significant source of information for at least a sub-set of investors with constrained attention, then we would expect its coverage decisions to have a direct influence on investors' knowledge and beliefs regarding the events in question. What if, instead, the WSJ's individual effect is negligible, but the financial news media as a whole represents a significant source of information for investors? In this case, if media coverage decisions exhibit commonality across reporters and news organizations, then, given its status as an industry leader, we might expect the WSJ to serve as a relatively good proxy for media coverage patterns more widely. Under both of these interpretations, we can think of the news media as informationally important due to its role of selecting and presenting particular information items for its readers, affecting investors' marginal costs of information acquisition and processing through their publication decisions.

Finally, imagine an even more restrictive case: realized coverage decisions in the financial news media do not directly impact investors' information sets at all. In this case, media coverage will, nonetheless, be useful as an indicator of investor attention to the extent that news outlets are able to model the information consumption preferences of their customers. In other words, if reporters are successful in applying a "theory of mind" regarding their consumers' underlying information preferences in order to predict which news items their readers will find most interesting (and, in a competitive news market, we might expect that they would be, on average), then observed coverage decisions will still serve as a guide for a researcher hoping to identify the characteristics of firms and events that will tend to attract attention more broadly.

This paper departs from previous studies in its identification of neglected firms based not only on firm characteristics, but also upon the very nature of the underlying information events that do (or do not) attract attention with respect to each firm. In this context, if attention allocation is asymmetric with respect to event characteristics, and if attention affects the speed with which stock prices respond to new information, this creates the potential for a directional drift in neglected firm values that might otherwise be indistinguishable from a more standard, symmetric risk-based story. In order to test this hypothesis, I disambiguate and compare the returns of neglected "good news" firms versus those of neglected "bad news" firms – if, as is suggested in the prior literature, the neglected firm premium is due to a symmetrically-distributed risk factor (e.g., information risk), then we should expect to observe excess return premia for both types of firms.

3 Data

This section describes the data set, consisting of information regarding the content of earnings announcements themselves, the characteristics of the firms making the announcements, as well as measures of related news coverage in the Wall Street Journal for the sample period from October 1984 to December 2005.

Earnings announcements and analyst data

The base event sample includes all earnings announcements that appear in I/B/E/S from 1984 to 2005 where 1) there are at lease two analyst forecasts in the previous month, and 2) the absolute surprise relative to the median analyst forecast is strictly less than the announcement day stock price. The resulting sample comprises 263,627 quarterly earnings announcement observations, including the announcement date, the announced normalized EPS number (i.e., expressed as a percentage of the announcement-day stock price), the number of distinct analyst EPS forecasts in the preceding month, the standard deviation of analysts' normalized EPS forecasts, the median analyst's normalized EPS forecast over the preceding month, and the quarter end-date to which the announced earnings number pertains.

Media coverage

I focus on media coverage in the Wall Street Journal (WSJ), which is considered to be one of the world's dominant financial news outlets, reaching approximately 2.6 million paying print and online subscribers with an average household net worth of US$ 2.5 million.[4] The media sample consists of 68,102 potentially-relevant Wall Street Journal articles from 1984 to 2005 (*Factiva Intelligent Indexing* code c151: "Earnings") in text format. In a two-stage classification strategy, articles are first categorized as either "earnings-related" or "not earnings-related", and then subsequently as "negative" or "not negative" (for simplicity, referred to as "positive" hereafter) using the "bag-of-words" computational linguistics tool *Rainbow* (McCallum (1996)).

Articles identified by the first-stage classification model as pertaining to a quarterly earnings announcement with posterior probability > 0.5 (49,113 articles) are then subjected to the second-stage classification. Those articles with a calculated posterior probability > 0.5 of being a member of the "negative" category are recorded as "negative" (17,280 articles), and all others are recorded as "not negative" (31,833 articles).[5] (For a more detailed description, please refer to Appendix A.)

With respect to each earnings announcement in the I/B/E/S event sample, I search among the media article observations for an identified earnings-related article within one week of the announcement date. If at least one identified earnings article pertaining to that firm is observed within the one-week window, I record that event as having received positive or negative coverage, depending upon how that article is classified. As we might expect, most firm-events are not observed to receive coverage in the WSJ: approximately 9% of announcements are associated with an identified positive earnings story, while approximately 4% can be linked with a related negative story.

Firm characteristics

Using data from COMPUSTAT, I calculate market to book ratios for each firm relative to each earnings announcement date by dividing the previous December 31 market value of equity by the book value of equity for the fiscal year ending in the prior year. Firms with M/B less than zero are discarded from the sample. I also consider the percentage of institutional ownership of equity observed as of the previous December 31. Stock return and trading volume data for the 60 trading days prior to each announcement window (i.e., days -61 to -2) are obtained from CRSP,

---

[4] http://www.dj.com/Products_Services/PrintPublishing/WSJ.htm

[5] While the "not negative" classification is referred to more simply as "positive" for the remainder of the paper, given the binary categorization scheme, the "not negative" category includes all earnings articles that would be classed as either "positive" or "neutral" under a ternary classification scheme. Given that double-negatives are relatively uncommon in standard English (e.g., one is unlikely to hear "Firm X *failed* to *disappoint* analysts' expectations" instead of "Firm X *exceeded* analysts' expectations"), this categorization method is expected to result in a somewhat sharper distinction between "positive" and "negative" overall semantic meanings. However, the inclusion of "neutral" articles under the *de facto* "positive" category implies that the "positive" vs. "negative" categorical distinction referred to herein should be interpreted in a relative rather than an absolute sense.

representing roughly the most recent quarter of daily data for each firm – average daily dollar-value of trading volume, average daily stock returns, and the standard deviation of daily returns are calculated for each of these pre-event windows. Stocks with a closing price less than $1 are discarded. Market size is shares outstanding multiplied by the stock price observed on day -2 relative to the announcement. Firms' industries are identified according to their "Fama-French 49" classifications (definitions available on Kenneth French's website). Data on institutional ownership are from Thomson Financial's CDA/Spectrum 13-F database.

After discarding firm-event observations due to missing values in CRSP, I/B/E/S, and/or COMPUSTAT, the final sample comprises 178,898 firm-events spanning a sample period from October 1984 to December 2005.

4 The Determinants of Media Coverage

In the first stage of the analysis, I investigate the market conditions and firm- and announcement-specific factors that predict absolute media coverage (i.e., an observation of either positive or negative coverage). I first consider factors based on firm characteristics. For example, we might predict larger firms to attract more attention – readers may be more interested in hearing about a firm with which they are already familiar, and a large firm will tend to have a greater number of people with a direct interest in the firms' prospects (e.g., employees, customers, suppliers, investors, etc.). Certain industries may also be favored, on average. Recent stock returns, analyst coverage, and trading activity are also considered.

A second set of potential predictors describe the information content of the earnings announcement itself. For example, if media attention is drawn to negative events (as we might expect to see if negative events are seen as relatively more sensational/interesting), then the announcement of an earnings number below analysts' expectations should result in a higher probability of coverage. Conversely, if media coverage is typically drawn to positive or "feel good" news events, then the observation of a loss should result in a lower probability of coverage, ceteris paribus. However, the appropriate specification to describe such a relationship is unclear. Is the size of the positive or negative surprise crucial, or does the media view such news in more categorical terms? I also investigate potential interactions – for example, might a loss combined with a negative surprise attract negative media coverage more reliably than the "sum of the parts"?

Finally, underlying economic and market-wide conditions may contribute to variation in coverage. I include recent returns on the S&P 500, and recent volatility in the S&P 500 to capture the potential effects of overall market conditions. For example, Veronesi (1999) presents a general equilibrium model wherein investors "overreact" to bad news in good times, and "underreact" to good news in bad times.

I apply a logit model to explain absolute coverage as follows:

$$\begin{aligned}
&Prob[ABS(COVERAGE)_{i,t} = 1] = F(SURPRISE_{i,t}, I_NEGSURPRISE_{i,t}, I_LOSS_{i,t},\\
&\quad I_NEGSURPISE_{i,t} \cdot I_LOSS_{i,t}, ANALYSTS_{i,t}, STDEV(FORECASTS)_{i,t}, B/M_{i,t},\\
&\quad RETURNS_{-1,i,t}, STDEV(RETURNS_{-1})_{i,t}, S\&P_{-1,i,t}, STDEV(S\&P_{-1})_{i,t}, MKTVALUE_{i,t},\\
&\quad USFIRM_{i,t}, \$VOLUME_{i,t}, INDUSTRY_{i}, YEAR_{i,t}, MONTH_{i,t}, DAY_{i,t}),
\end{aligned}$$

where $ABS(COVERAGE)_{i,t}$ equals one if there is an identified WSJ article associated with the announcement, and zero otherwise; $SURPRISE_{i,t}$ is the announced EPS minus the median analyst forecast from the prior thirty days, divided by the stock price; $I\_NEGSURPRISE_{i,t}$ is a dummy variable equal to one if $SURPRISE_{i,t}$ is negative, and zero otherwise; $I\_LOSS_{i,t}$ is a dummy variable equal to one if announced EPS is negative, and zero otherwise; $ANALYSTS_{i,t}$ is the natural log of the number of distinct analyst forecasts observed in I/B/E/S in the 30 days preceding the announcement; $STDEV(FORECASTS)_{i,t}$ is the standard deviation of the normalized analyst EPS forecasts recorded in I/B/E/S in the 30 days preceding the announcement; $MKTVALUE_{i,t}$ is the natural log of the firm's market value of equity prior to the event window; $RETURNS_{i,t}$ is the firm's cumulative stock return over the 60 trading days prior to the event window; $I\_USFIRM_{i,t}$ is a dummy variable equal to one if the firm is identified as a U.S. firm in CRSP; $\$VOLUME_{i,t}$ is the natural log of the average value of daily stock trading over the 60 days preceding the announcement window; $INDUSTRY_i$ represents a set of dummy variables for the Fama-French 49 industry classification; $YEAR_{i,t}$ represents the set of year dummy variables; and $MONTH_{i,t}$ and $DAY_{i,t}$ represent month-of-the-year and day-of-the-week dummy variables to control for potential seasonal effects (see, e.g., DellaVigna and Pollet (2008)).

Table 2 presents the results of a logit regression of $ABS(COVERAGE)_{i,t}$ on market conditions, firm and event characteristics, and the time and industry dummies. With respect to the event-specific determinants of unsigned media coverage, I find that losses and negative earnings surprises (relative to median analyst expectations) are more likely to attract media coverage. As mentioned earlier, this may be related to, for example, the media's often-cited propensity to focus on "sensational" and/or unexpected stories. Table 1 shows that accounting losses are relatively infrequently observed, so there may tend to be a certain degree of "surprise" attached to each such announcement. In particular, a firm typically attempts to present a positive (or, at the very least, an ambiguous) picture of its performance to the market. Thus, an unambiguously negative piece of news might naturally be framed as "surprising" relative to the communication lines that firms generally attempt to put forward. Note, however, that this increased probability of coverage regarding bad news is relative: negative news articles are still less frequently observed than non-negative ones in absolute terms (i.e., 4% versus 9%). Not surprisingly, I also find that large firms and those with high analyst coverage and high trading volumes are most likely to receive attention in the media.

Figure 1 illustrates the relationships between firm and event characteristics and probability of coverage as described in Table 2. Holding event characteristics constant, any news regarding large, important firms is more likely to receive coverage. On the other hand, holding firm characteristics constant, a negative event is more likely to receive media coverage than a positive one. The implication is that negative news events regarding large firms are most likely to receive media coverage, while positive news events involving small firms are least likely to receive coverage. In short, the results in Table 2 support the asymmetric attention hypothesis – this is a crucial finding for the analysis that follows.

Having examined the determinants of coverage in an absolute sense, I make use of the positive/negative categorical distinction. For example, the observation of a loss and/or a negative surprise might be expected to predict a negative article.

With respect to positive or negative coverage, I test the following relationship as a multinomial logit:

$$\begin{aligned}\text{Prob}[\text{COVERAGE}_{i,t} = j] = G_j(&\text{SURPRISE}_{i,t}, \text{I\_NEGSURPRISE}_{i,t}, \text{I\_LOSS}_{i,t},\\ &\text{I\_NEGSURPISE}\cdot\text{I\_LOSS}_{i,t}, \text{ANALYSTS}_{i,t}, \text{STDEV(FORECASTS)}_{i,t}, \text{B/M}_{i,t},\\ &\text{RETURNS}_{-1,i,t}, \text{STDEV(RETURNS}_{-1})_{i,t}, \text{S\&P}_{-1,i,t}, \text{STDEV(S\&P}_{-1})_{i,t}, \text{MKTVALUE}_{i,t},\\ &\text{USFIRM}_{i,t}, \text{\$VOLUME}_{i,t}, \text{INDUSTRY}_{i},\text{YEAR}_{i,t}, \text{MONTH}_{i,t}, \text{DAY}_{i,t}),\\ &\qquad\qquad j= -1, 0, 1\end{aligned}$$

where $\text{COVERAGE}_{i,t}$ is equal to 1 if there is an identified positive earnings story related to the announcement, equal to -1 if there is an identified negative earnings story related to the announcement, and 0 otherwise.

Table 3 presents the results from the multinomial logit regressions. As expected, positive event characteristics (such as positive surprises relative to analysts' expectations) predict positive media coverage, and *vice versa*. Firm characteristics that would be expected to attract attention unambiguously, such as firm size, analyst coverage, and average daily dollar value of stock trading, likewise behave as expected. Somewhat more interesting is the observation that some firm characteristics appear to have asymmetric effects with respect to predicting positive or negative media coverage. For example, high analyst forecast dispersion predicts a greater likelihood of negative coverage but a lesser likelihood of positive coverage, perhaps reflecting the impact of increased uncertainty. Similarly, high realized stock volatility in the period preceding the announcement predicts negative coverage, and implies a smaller probability of positive coverage. Somewhat counter-intuitively, high recent stock returns appear to predict smaller probabilities of both positive and negative coverage. Pre-event market conditions also seem to predict coverage: recent S&P volatility seems to predict positive coverage while high recent S&P returns predict negative coverage.

It should be noted that the foregoing results on the determinants of coverage highlight a potential problem with attempting to estimate the impact of media coverage on contemporaneously-observed returns in an event study setting: many of the event characteristics that predict positive and negative coverage may also affect announcement returns. In other words, there is a very real possibility that an apparent relationship between, e.g., negative media coverage and negative stock returns could be due not to the impact of coverage, but, rather, to the fact that reporters tend to write negative stories about firms that report losses, fail to meet analysts' expectations, etc.

Discussion of results

One explanation regarding reader interest simply recognizes that news organizations are motivated to sell news, while news consumers are more likely to be interested in reading about events related to firms with which they are already familiar or have some personal interest. Firm size can be seen as a proxy for the number of customers, employees, suppliers, etc., that a firm possesses. Similarly, the average dollar value of recent stock trading should be related to the number of active investors in a stock and the intensity of their interest. Therefore, it is no surprise that these proxies are positively related to probability of coverage. Fang and Peress (2007) examine a comprehensive set of firm-related media stories which they attempt to match to the entire universe of stocks. Since they observe both when firms do and do not receive media coverage, they are able to identify firm characteristics that predict media coverage. In particular, they find that size, B/M, analyst coverage, and individual ownership are all associated with a higher probability of observing some news story with respect to that firm. At the same time, since their measure of coverage is not centered on identification of comparable events across firms, it is difficult to identify the event-specific (as opposed to firm-specific) characteristics that tend to attract attention. Engelberg (2007) presents similar evidence with respect to size and analyst coverage. Analogously, Kaniel et al. (2007) find that total net assets under management are positively related to the probability of media coverage for mutual funds.

With respect to the event-specific determinants of coverage, while a finding of negativity bias in media attention is consistent with Soroka (2006) and Harrington (1989)[6], it may be somewhat puzzling in light of results from Kaniel et al. (2007) and Engelberg (2007). For example, Kaniel et al. (2007) find that positive recent performance predicts higher levels of media coverage for mutual funds. What might account for this apparent difference in the event-specific determinants of coverage? While it is difficult to speculate, one potential explanation has to do with endogeneity in the production of the information used by reporters. If mutual funds are more likely to produce information (e.g., issue press releases) when performance has been good, and if independent sources of information regarding fund performance are relatively inaccessible, then this may naturally result in higher levels of coverage for these funds. Similarly, Engelberg (2007) presents mixed evidence that the probability of observing media coverage is positively related to contemporaneous event-window stock returns; in addition to the potential issue of simultaneity, the media coverage data in this setting (i.e., DJNS) may be more likely to include firm-originated reports, so it is possible that self-selection of positive news stories may be a factor here as well. However, the explanation for these seemingly anomalous results remains unclear – this is an area for further study.

5 Probability of Media Coverage (PMC)

In this section, I describe the construction of the PMC measures and discuss their distributional characteristics and statistical relationships with respect to other variables of interest.

Construction of PMC

Positive, negative, and absolute probability of media coverage with respect to each firm-event is calculated as the predicted probability of positive, negative, or any (i.e., positive or negative) coverage, respectively, from the multinomial logit regression specified in Table 3, column 5 (omitting the year dummies).

$$\begin{aligned} PMC_{i,t} &= \text{Prob}\{\text{Positive or negative media coverage} \mid X_{i,t}\} \\ &= 1 - \text{Prob}\{COVERAGE_{i,t} = 0 \mid X_{i,t}\} \\ &= PMC^{+}_{i,t} + PMC^{-}_{i,t}, \end{aligned}$$

$$PMC^{+}_{i,t} = \text{Prob}\{\text{Positive media coverage} \mid X_{i,t}\}$$

[6] The finding of negativity bias is also consistent, albeit somewhat indirectly, with broader evidence regarding asymmetrically small reactions to positive news and/or large reactions to negative news in, e.g., McQueen et al. (1996), Veronesi (1999), Kothari et al. (2005), Skinner and Sloan (2002), and Conrad et al. (2002).

$$= \text{Prob}\{\text{COVERAGE}_{i,t} = 1 \mid X_{i,t}\}$$

$$\text{PMC}^{-}_{i,t} = \text{Prob}\{\text{Negative media coverage} \mid X_{i,t}\}$$
$$= \text{Prob}\{\text{COVERAGE}_{i,t} = -1 \mid X_{i,t}\}$$

where $X_{i,t}$ includes the identified firm and event characteristics that were found earlier to predict media coverage.

Figure 2 illustrates that predicted attention probabilities for most firm events lie relatively close to zero – this is as expected, given that the unconditional probability of WSJ coverage in the sample is 13.1%. Simply, there are typically many more earnings announcements made during a given day or week than could be the subject of stand-alone articles in the WSJ, even if all of them were thought to be potentially "interesting enough" to warrant such coverage. As the results in Tables 2 and 3 demonstrate, however, the observed high degree of selectivity in coverage decisions here does not appear to present a problem for the statistical identification of the determinants of coverage – the sample is more than large enough to accommodate a high percentage of zeros on the LHS. At the same time, the highly non-normal distribution of PMC might cause us to hesitate before using it as a raw input in further analysis. With this in mind, the relative rankings provided by PMC, rather than the raw values themselves, is the primary focus in the analysis that follows.

PMC and firm and event characteristics

Panel A of Table 4 compares average firm and event characteristics by absolute probability of media coverage decile. Given the results in Table 2, it is not surprising that low-attention firms are typically smaller firms with less analyst activity, smaller trading volumes, higher recent stock returns, and positive earnings surprises.

Panels B and C look at characteristics across positive and negative PMC deciles, respectively. Firms in the smallest PMC^{+} decile (i.e., those least likely to receive positive news coverage) are relatively small, low-trading volume, low analyst coverage, low B/M, low institutional ownership firms that have experienced relatively negative earnings news and have low and stable recent stock returns. Firm in the smallest PMC^{-} decile (i.e., those least likely to receive negative news coverage), on the other hand, tend to be largely similar in profile, except with relatively positive earnings news and higher, more volatile recent returns.

Surprisingly, while we might have expected relatively low-attention events to elicit relatively weak event-window price reactions, despite the typically positive tenor of such news, this does not seem to be the case. Comparing the final row of column 1 in panel A to the final row of column 10 in panel B, we observe that the average announcement return for firms in the lowest PMC^- decile is actually higher than the average announcement window return for firms in the highest PMC^+ decile. While the difference in average event-window returns is relatively modest, and we do not account here for other firm and event characteristics that may affect, for example, the expected volatility of short-term announcement-window returns, it is nonetheless clear that the typically good news in low-PMC and low-PMC^- events is indeed "noticed" by market participants despite the very low probability of media coverage.

PMC and industry classification

Figure 3 presents average predicted probability of media coverage by Fama-French 49 industry classification for the top-15 and bottom-15 industries (attention probabilities calculated from the regression in the last column in Table 3), illustrating that firms in some industries seem to be unconditionally more likely to receive coverage than others. A casual inspection suggests that abnormally high propensities for media coverage seem to be associated with "high-profile" industries that might typically possess large customer and/or employee populations, or higher advertising expenditures, etc. (e.g., Tobacco Products, Beer and Liquor, Printing and Publishing, Recreation, Apparel, Automobiles and Trucks, and Retail) – this makes sense if we consider that reporters may prefer to report on firms with which the average reader is more likely to already be familiar; this observation seems to generally confirm the intuition discussed earlier with respect to expected reader interest and firm size, etc. Other, seemingly lower-profile, less-glamorous industries (such as Electrical Equipment, Shipping Containers, Medical Equipment, and Measuring and Control Equipment, etc.) seem to be over-represented among those industries with low average probability of media coverage. It is interesting to note also, that, while PMC^+ seems to increase mostly monotonically with absolute PMC, PMC^- does not seem to do so. This may be due to the fact that PMC^- seems to be somewhat more sensitive to event-specific (e.g., good news vs. bad news) rather than certain firm-specific (e.g., large vs. small) characteristics, which are likely to be more homogenously distributed within industry classes.

6 Attention and Neglect in the Cross-Section of Stock Returns

In this section, I explore the relationship between investor attention and expected returns. First, I do this by examining the time-series of returns on portfolios based on the PMC measures, controlling for commonly-identified risk factors. Second, I perform pooled cross-sectional regressions of individual returns on PMC and firm characteristics. Third, I examine the time series of monthly cross-sectional regressions on PMC decile membership in Fama-MacBeth tests.

Portfolio formation

The monthly portfolios are formed by sorting all of the stocks by their most recent PMC, PMC^{+}, and PMC^{-} observations within the prior three months. For example, assume that firm A makes an earnings announcement on February 15th, with firm and event characteristics resulting in media coverage probabilities of PMC = 0.78, PMC^{+} = 0.22, and PMC^{-} = 0.56. Deciles are formed at the beginning of each month with respect to each set of scores. Firm A will then be included in the PMC decile portfolios formed on March 1, April 1, and May 1 based on the Feb. 15th observation. If A releases its next earnings announcement on April 23rd, then its May 1 decile assignments will instead be based on those more recent predicted probability values, etc. However, if A does not report its next set of results until June 4th, then it is dropped from inclusion in the PMC portfolios formed at the beginning of June, before returning again in the July, August, and September portfolio assignments. Portfolio returns are equal-weighted and are based on the closing prices observed on the last trading day of the month.

Table 4 presents average firm and event characteristics by PMC deciles. Panel A shows the results for firms sorted by absolute PMC (i.e., probability of positive or negative coverage). Looking at the first column, we see that firm-events with the lowest predicted probability of receiving coverage are typically small, low-trading volume, low-analyst activity, low-beta, low-institutional ownership firms with better-than-expected earnings, positive and relatively volatile recent stock returns, and positive announcement-window CARs[7]. On the other hand, firm-events with the highest probability of coverage are typically the converse: larger, high-trading volume, etc., firms with lower-than-expected earnings results, and smaller (although still positive) recent stock returns and announcement-window CARs.

Panels B and C differentiate media attention into positive and negative coverage. Looking at the first columns of the two panels, we see that smaller firms with low analyst activity, etc., are least likely to attract both positive and negative coverage. However, the event-specific and recent stock performance factors show some striking differences, as we might expect. Firm-events least likely to attract positive media coverage are those with respect to firms that failed to meet analyst expectations on average, with relatively low recent stock returns and negative announcement returns. On the other hand, firm-events least likely to attract negative media coverage typically

[7] Where $CAR_{i,t} = [R_{i,-1} - R^e_{-1}] + [R_{i,0} - R^e_0] + [R_{i,+1} - R^e_{+1}] + [R_{i,+2} - R^e_{+2}]$; $R_{i,t}$ is stock i's return on day t (relative to the announcement date) and R^e_t is the expected return calculated using coefficients Carhart (1997) 4-factor model estimated on the 60 trading days prior to the event window.

beat analysts' expectations and experienced positive announcement returns. Again, these results are as expected, given the results in Table 3.

Media coverage and expected returns: portfolio tests

I examine excess returns on the monthly PMC decile portfolios. Accounting for sources of return previously identified in the literature may be particularly important because we know that the PMC measures load heavily on several firm and event characteristics that are analogous to risk-factors such as size, value, momentum, etc. If, for example, significant excess returns on an PMC portfolio were observed to disappear when the Fama-French (1993) factors are added to the return model, this might indicate that any apparent outperformance was actually due to the portfolio being heavily loaded with, e.g., small-sized and/or high book-to-market stocks. I begin with the CAPM model:

$$R_t = \alpha + \beta MKT_t + e_t \qquad\qquad t = 1,...,T \ ,$$

where R_t is the monthly return on an equal-weighted PMC portfolio less the risk-free rate and MKT_t is the market return minus the risk-free rate. Panel A of Table 5 presents estimates of α for each of the PMC decile portfolios; in the final column, I show the results for a zero-investment portfolio that is long stocks in the smallest decile (i.e., firms least likely to attract absolute, positive, or negative media coverage) and short stocks in the largest decile (i.e., firms most likely to attract absolute, positive, or negative media coverage). Looking at the first row of panel A, we see that the portfolio that is long low-absolute attention stocks and short high-absolute attention stocks yields returns of almost 85 bps per month, results which are both statistically and economically significant. Differentiating between positive and negative coverage, the estimated alpha for the long-short PMC^+ portfolio is not significant, while the long-short PMC^- portfolio yields excess returns of approximately 154 bps per month, or over 20% per year.

Proceeding to a less parsimonious specification, panel B presents estimates from a Fama-French 3-factor model:

$$R_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + e_t \qquad\qquad t = 1,...,T \ ,$$

where R_t and MKT_t are as defined above, and SMB_t and HML_t are the size and value risk factors available on CRSP. Here, the estimated alphas on the long-short PMC and PMC^- portfolios fall slightly, to approximately 70 bps and 123 bps per month (8.7% and 15.8% per year) respectively.

Since we have noted that the zero-investment PMC and PMC^- portfolios are effectively long stocks with relatively good recent performance and short stocks with relatively poor recent performance (both in terms of the information in the earnings announcement itself and in terms of stock returns prior to the announcement), it may be crucially important to account for price momentum, which has previously been shown to explain significant return predictability in the cross-section. Panel C presents estimates from the following regression:

$$R_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 WML_t + e_t \qquad t = 1,...,T \ ,$$

where R_t, MKT_t, SMB_t, and HML_t are as described earlier and WML_t is the "winners minus losers" momentum factor from CRSP. Interestingly, while the alpha estimate on the long-short PMC$^-$ portfolio drops to 75 bps per month (9.4% per year), excess returns on the long-short PMC portfolio remained unchanged from the 3-factor model above.

Finally, researchers have identified liquidity as an important potential risk factor explaining the cross-section of stock returns (e.g., Pastor and Stambaugh (2003)). Since low-PMC firms tend to be smaller firms with lower trading volumes, it is important to check whether the apparent excess returns might actually be due to a liquidity premium. Panel D presents estimated alphas from the following model.

$$R_t = \alpha + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 WML_t + \beta_5 PS_t + e_t \qquad t = 1,...,T \ ,$$

where R_t, MKT_t, SMB_t, HML_t, and MOM_t are as defined above and PS_t is either the Pastor-Stambaugh (2003) liquidity factor (in levels) available from CRSP.[8] The addition of the liquidity factor lowers the estimated alphas by a couple of basis points, but the results are basically unchanged.

The findings reported above are consistent with Fang and Peress (2007), where they find that firms with no observed coverage subsequently outperform those with high coverage. But where might the outperformance be coming from? The results in the second and third rows of panels A, B, C, and D, where we are now able to make a distinction between positive and negative news events, give us a clear answer. While the long-short PMC$^+$ portfolios generate returns that are not significantly different from zero, the long-short PMC$^-$ portfolio generates returns from a high of 154 bps per month in the CAPM model to a low of 74 bps per month in the five-factor model. In the previous section, I presented evidence for asymmetric attention with respect to positive and negative news. Here, we find support for the hypothesis that the systematic inattention to neglected "good news" firms in the information environment has strong implications for pricing: after controlling for all of the risk factors commonly cited in the literature, the long-short PMC and PMC$^-$ portfolios yield returns that are potentially both economically and statistically significant. On the other hand, the portfolio strategy focusing on neglected "bad news" firms does not yield the significant negative abnormal returns that we

[8] In untabulated results, I test alternate specifications with the Pastor-Stambaugh "innovations" version as well as Sadka's (2006) liquidity factors – the results are similar.

would expect to see if the low-attention effect was symmetric. In short, the results here are more consistent with systematic underreaction to particular public news events, as in, e.g., Hong, Lim, and Stein (2000).

Factor loadings

Table 6 shows 4-factor loadings for the PMC portfolios.[9] In the final column of panel A, note that the long-short portfolio formed on absolute PMC (i.e., long firms with low expected probability of positive or negative coverage, and short firms with high expected probability of coverage) loads positively on HML and SMB. Given that relatively large, relatively high book-to-market firms are most likely to attract media coverage (Table 2 and Table 4), this is not surprising. Similarly, we saw earlier that high beta firms attract attention, so the long-short portfolio loads negatively on the market return. Interestingly, all of the PMC decile portfolios load negatively on WML, albeit with relatively small coefficients.

In panel B, we begin to see some clear differences among portfolios formed on absolute, positive, and negative attention probabilities. In contrast to the results above, the long-short PMC^+ portfolio loads negatively on HML and negatively on WML. In particular, looking at the first column of panel B, firms with the lowest probability of receiving positive coverage (i.e., the "low-profile, bad news" firms) have much smaller coefficients on HML and WML than we saw for the firms with the lowest probability of absolute coverage (the first column of panel A). Note the negative coefficient WML; this is as we might expect: by sorting on positive attention, the long-short PMC^+ portfolio is long stocks with relatively bad news in the previous month and short stocks with relatively good news.

Panel C shows the results for portfolios formed on PMC^- (expected probability of receiving negative attention). In the final column, we see that the long-short portfolio now loads positively on WML, as the portfolios with the highest probability of negative coverage load increasingly negatively on this factor. Otherwise, the factor loadings here are identical in sign to those with respect to the absolute PMC long-short portfolio in panel A.

Double-sorts by firm and event characteristics

Since the PMC measures load on a number of variables that we might otherwise think of as potentially affecting expected returns, I examine the returns of long-short PMC portfolios formed within firm- and event-characteristic quintiles. In particular, while we know that the portfolio strategies appear to yield positive excess returns in the aggregate, it is important to find out where

[9] The Pastor-Stambaugh liquidity factor is omitted from the specifications in Table 6 due to statistical insignificance – results available upon request.

the trading profits are coming from. I concentrate here on returns from portfolios based on quintiles of PMC^-. Panel A of table 7 presents the results sorted first on market value. Interestingly, while the long-short PMC^- portfolio yields significant positive excess returns only in the bottom three quintiles of market value, returns are actually highest within the middle quintile.

Panel B looks at portfolios formed within analyst coverage quintiles. Here, while abnormal returns appear to be concentrated in the lower half in terms of analyst activity, profits are highest among trades within the lowest quintile. Panel C shows the results for portfolios sorted first on book-to-market of equity. Excess returns are highest for the high B/M firms, but they are not nearly as concentrated as in the two earlier cases – even within the lowest B/M quintile, we observe significant positive returns (albeit only significant at 10%). In panel D, we see that profits are highest among low-beta firms, although, again, excess returns remain positive and (modestly) significant up to the largest quintile. Finally, panel E presents the results from sorting on recent returns: here, the long-short portfolios on PMC^- are most profitable among firms with positive recent returns – this is not entirely surprising since high recent stock returns is presumably one of the most important "good news" elements that is being ignored for firms with low PMC^-.

Earnings momentum and media coverage

The profitable portfolio trading strategies described above essentially prescribe going long a set of stocks that, on average, have experienced "good earnings news" in the previous 3 months. This is a consequence of the fact that low attention probabilities are explained, in part, by the observation of positive profits and positive earnings surprises (see Tables 2 and 3). This inevitably leads to the question: Is what we are seeing simply post-earnings announcement drift (PEAD)? After all, earnings momentum strategies advise going long stocks with large positive earnings surprises and shorting stocks with large negative surprises. However, while there is obviously an element of PEAD at work here, there are two lines of evidence against a "standard", symmetric PEAD story.

First, a traditional PEAD explanation would imply that the part of the portfolio that is short negative earnings surprise stocks should also be profitable, as well as the one that is long positive surprise stocks. As we see in Table 4, the stocks in the largest PMC and PMC^- deciles have experienced negative earnings surprises on average. However, referring to Table 5, we observe that the excess returns on these largest decile portfolios are never significantly negative. Finally, Hou et al. (2006) observe that earnings momentum profits are highest for low-attention stocks – however, note here that returns on the long-short PMC^+ portfolio (which is long the relatively low-attention, bad-news stocks) are not significantly different from zero, implying that the underreaction is asymmetric with respect to good and bad news.

Second, symmetric PEAD would predict that the PMC-based strategies should work equally well across earnings surprise quintiles – that is to say, not all that well, since forming portfolios within earnings surprise quintiles will reduce the differences among the stocks. For example, if earnings momentum were behind the results, by forming PMC portfolios within the largest earnings surprise quintile we would effectively be going long a portfolio of firms with "best of the best" surprises and going short a portfolio with "worst of the best" surprises. Obviously, we would not expect this to be as profitable as a strategy that is long the "best of the best" surprises and short the "worst of the worst". Looking at the results in Table 7, panel B, however, it is clear that the PMC- and PMC^--based strategies are actually *more* profitable within the largest earnings surprise quintile than they were with respect to the sample as a whole (Table 5, panel C). Alphas from the 3rd and 4th earnings-surprise quintiles are smaller, but also positive and significant, before falling to insignificance in the two smallest quintiles.

Cross-sectional analysis

In order to test for the potential impact of PMC in the cross-section, I perform the following regression:

$$R_{i,t} = \phi_0 + \phi_1 \log(PMC)_{i,t-1} + \theta Controls_{i,t} + \varepsilon_t \qquad t = 1,...,T\ ;\ i = 1,...,N\ ,$$

where $R_{i,t}$ is the month *t* return for stock *i,* $\log(PMC)_{i,t-1}$ is the lagged natural logarithm of a PMC measure, and $Controls_i$ includes a set of characteristics that may help to explain returns (including $B/M_{i,t}$, $MKTVALUE_{i,t}$, $STDEV(RETURNS_{-1})_{i,t}$, etc., as defined earlier). Table 8 shows the results with standard errors robust to arbitrary heteroskedasticity and clustering by firm.[10] In the first three columns, monthly returns are regressed on lagged log(PMC), log(PMC^+), and log(PMC^-) in turn, along with a number of "standard" controls. Consistent with the earlier portfolio results, I find that low PMC and low PMC^- predict higher subsequent returns. However, in the cross-section, low PMC^+ now appears to predict lower returns. While the size of the estimated coefficient on log(PMC^+) suggests that the effect is not quite as large as with log(PMC) and log(PMC^-), this finding of a significant positive coefficient seems to contradict the earlier portfolio results. While this result goes against an information risk or illiquidity-based explanation (i.e., a low expectation of positive attention should unambiguously predict higher returns if "information risk" is priced and media coverage is a proxy for such risk), there are at least two potential explanations for this finding. The first is that low-attention, bad news firms

[10] Results are similar with clustering by year or industry, or with firm fixed effects.

(i.e., firms with the lowest probability of attracting positive attention) subsequently experience negative returns as ignored information is slowly impounded into prices. If true, this would imply that low-attention drift is somewhat more symmetric than we would have thought based on the portfolio results.

A second potential explanation for the positive coefficient on log(PMC^+) is that we may be seeing residual post-earnings announcement drift: firms with the highest probability of positive coverage (who, in general, will have experienced positive earnings surprises) experience positive returns in the following month, and the converse with respect to firms with the lowest probability of positive coverage.[11] In columns 4 to 6 of Table 8, I include the earnings surprise (among other controls) as an explanatory variable. In this specification, the coefficient on log(PMC^+) falls (although we still cannot reject significance at 10%), while the coefficient on $SURPRISE_{i,t}$ is significant and positive. On the other hand, the coefficients on log(PMC) and log(PMC^-) in columns 4 and 6 are observed to grow larger in absolute magnitude.

Fama-MacBeth regressions

I perform Fama-MacBeth tests by regressing stock returns on PMC variables within each month and then examining the resulting time series of monthly estimated coefficients. Panel A of Table 9 presents the results from regressing returns on dummy variables indicating membership in the first or tenth deciles of PMC for the previous month.

$$R_{i,t} = \phi_0 + \phi_1 I\{PMCdecile\}_{i,t-1} + \theta Controls_{i,t} + \varepsilon_t \qquad t = 1,...,T\ ;\ i = 1,...,N\ ,$$

where $I\{PMCdecile\}_{i,t}$ is equal to one if firm i was in the first or tenth decile of PMC, PMC^+, or PMC^- for the previous month, and zero otherwise; and $Controls_{i,t}$ contains the set of control variables in the first specification from Table 8.

Note that the average coefficients on the first decile PMC and PMC^- dummies are significant and positive, while the average coefficient on the tenth decile PMC^- dummy is significant and negative. The positive coefficients on the first decile dummies are as we might expect, given the preceding results for both the time-series tests and the pooled cross-section (i.e., low-attention, good-news firms experience high subsequent returns). The result for the tenth decile PMC^- dummy is somewhat more surprising, suggesting that bad-news high-attention firms experience downward drift.

Instead of using the binary measures, Panel B of Table 9 presents the results of regressing returns on the natural log of the lagged PMC variables.

[11] Furthermore, if PEAD is having an effect on the results, the estimated coefficients on PMC and PMC^- may be too small (since PMA and PMA^- are negatively related to earnings surprise).

$$R_{i,t} = \phi_0 + \phi_1 \log(PMC)_{i,t-1} + \theta Controls_{i,t} + \varepsilon_t \qquad t = 1,...,T \; ; \; i = 1,...,N \; ,$$

where the variables are as defined earlier.

Sub-period results

Are the findings sensitive to the sample period? The Fama-MacBeth results would seem to indicate that the estimates are fairly consistent over time, but we may be interested in finding out how simulated portfolio returns might have changed. Is this a strategy that might still work today? In untabulated results, I divide the sample into four sub-periods (1984-1989, 1990-1994, 1995-1999, 2000-2005), and then re-run the four-factor portfolio tests with respect to each one. The estimated monthly alphas for the long-short PMC^- portfolios are[12]: 188 bps*** for 1984-1989; 64 bps** for 1990-1994; 39 bps (no statistical significance) for 1995-1999; and 125 bps*** for 2000-2005. In the late-90s sub-period, the culprit appears to be 1999 – when this year is omitted, the estimated portfolio alpha for the remaining 1995-1998 period is 105 bps*** per month.

6.2 Neglected Firm Effect or Delayed Response to Positive News?

The foregoing results indicate that PMC has significant incremental predictive power with respect to stock returns. Estimated excess annualized PMC^- portfolio returns range from a high of 20% in a CAPM model to just over 9% in a five-factor model that includes both momentum and liquidity factors. These findings are consistent with (albeit significantly larger in magnitude than) previous research finding that low-media coverage predicts high returns. Fang and Peress (2007) identify a return premium of over 3 percent per annum on stocks that are observed not to receive media coverage versus those that receive high attention; they show that this effect cannot easily be explained by the standard set of risk factors. In a similar vein, Gadarowski (2002) finds that high news coverage predicts lower subsequent stock returns.

More broadly, looking beyond the potential impact of media coverage in particular, researchers have long sought to explain the puzzles of apparently high risk-adjusted returns for groups of firms which could be characterized as possessing relatively severe frictions or risks in their information environments, e.g., with respect to small firms (e.g., Banz (1979), and Reinganum, (1981)), firms that are "neglected" in terms of analyst coverage (e.g., Arbel and Strebel, (1982)), and "price delayed" firms (Hou and Moskowitz, (2005)). Explanations have typically focused on liquidity (Amihud and Mendelson (1986)), or on potential frictions and risks

[12] * significant at 10%; ** significant at 5%;*** significant at 1%

in the information environments faced by these firms (e.g., Merton (1987), or Easley et al. (2002)).

Easley et al. (2002) find that probability of informed trading (PIN), which they interpret as a proxy for information risk, has predictive power with respect to the cross-section of returns. In terms of media coverage, one potential explanation for my low-attention premium is that these firms are subject to higher levels of information risk.

Sadka (2006) argues that a substantial proportion of both momentum and PEAD returns can be explained by liquidity risk. Again, I do not find here that any of the liquidity factors examined are significant in explaining PMC portfolio returns. In addition, while it is true that low-attention firms have many of the characteristics that we would typically associate with low liquidity, the liquidity explanation is hard to reconcile with the observation that only the low-attention good news firms seem to trade at a discount.

At a very basic level, this paper presents evidence of a return premium for firms that are likely to be neglected in terms of media attention. How, then, do my results differ from the aforementioned papers, in particular Fang and Peress (2007)? In this study, I construct an event-specific measure of neglect, documenting significant asymmetry in financial news media coverage decisions with respect to the content of the underlying information shock. Making use of this observation, I identify and investigate a new explanation for the finding of higher performance for low-attention (neglected) stocks in this context: asymmetric underreaction with respect to positive news for these firms.

My results support the asymmetric underreaction hypothesis, while they do not tend to support the information risk and investor recognition explanations. Specifically, when I look at propensities to attract positive and negative coverage separately, I do not find evidence that the relatively low-attention, bad news firms (i.e., firms with the lowest probability of positive coverage) experience either significantly positive excess returns (as we would expect to see if these stocks were being discounted due to information risk) or significantly negative returns (as we would expect if the low-attention effect was symmetric with respect to positive and negative news events).

While the observation of apparent underreaction to positive news is consistent with most of the prior research mentioned, it would seem to be rather at odds with the findings of HLS, who argue that "bad news travels slowly". The authors find that momentum strategies are more profitable for relatively small firms and those with low analyst coverage; furthermore, they show that the effect is more pronounced for recent losers than recent winners, which they interpret as evidence that, given managers' reluctance to communicate bad news, the presence of analysts is more important for "drawing out" negative information. One important point is that the authors' explanation is potentially consistent with a negativity bias in media coverage: if bad news is more likely to be credible/truthful, then it may be natural for information intermediaries to feature it more prominently once such an item has already been "drawn out" or made public (as in the case of earnings announcements). Also, note that the firms and the information flows being considered are potentially quite different. In this paper, I investigate potential underreaction to a particular set of underlying public events that are characterized by the fact that they are unlikely to have had a large, immediate impact upon prices. In contrast, HLS focus on drift following unspecified information flows which are actually identified (albeit implicitly) by the observation of a significant contemporaneous price movement. Finally, HLS identify their strongest evidence of asymmetric price momentum with respect to a subset of firms that are generally much smaller and may have significantly worse information environments than those included in my sample. In particular, while all of the firms considered in this study have at least two distinct analyst forecasts in the month prior to their announcements, HLS include all firms above the 20^{th} percentile NYSE/AMEX size breakpoints, a significant proportion of whom (e.g., 41.7% in 1988 for the 20^{th}-40^{th} percentile by size) are observed to have no analyst coverage.

7 Robustness Tests

The foregoing results showing high excess returns from trading strategies based on predicted media coverage, (and, in particular, the evidence of asymmetry with respect to positive and negative information), are certainly suggestive of a link between investors' attention allocations and the market's underreaction/overreaction to news events, but we are left with several unanswered questions. In particular, what if the simulated low-PMC trading strategies are simply being rewarded for holding stocks that are, e.g., small and/or have high market betas, etc.? It is clear that PMC's ability to predict returns, by its very construction, must derive from the identified determinants of coverage in some sense. Since PMC is the estimated probability resulting from a regression of observed coverage, PMC is essentially a non-linear combination of the regressors (recall that this is one of the key advantages of this measure, in that it allows us to side-step some potential problems related to omitted variable bias and measurement error). We might hope that the four-factor setting would correct for such effects in portfolio returns with respect to these standard risk factors, but there are no guarantees that this is sufficient. In short, it is possible that PMC may simply identifying stocks using an optimal weighting of previously-identified risk factors. In particular, market size and past returns were found to be highly significant predictors of coverage. Furthermore, other variables, such as analyst attention, may also be highly correlated with size, etc.

In this section, I address this potential concern by examining the predictive content of restricted versions of the expected coverage measures. Specifically, I examine PMC variants projected on different sets of predictors in turn, excluding size, B/M, momentum, and beta in all specifications. In order to eliminate any lingering effects from covariation with the remaining regressors, I then orthogonalize each restricted measure by taking the residual from a regression of PMC on size, B/M, momentum, and beta. Decile portfolios are then formed with respect to each measure. Table 10 presents the four-factor portfolio alphas based on these restricted and orthogonalized versions of the base PMC measures.

Panel A shows the results for portfolios formed on PMC1, a specification that only includes information on the earnings release itself. The results here are consistent with a generalized underreaction to positive earnings news: both the highest decile of $PMC1^+$ and the lowest decile of $PMC1^-$ (i.e., the "good news" stocks) have high predicted returns. Since positive coverage is unconditionally more likely to be observed than negative coverage, and we are controlling for no other factors, absolute PMC is positively related to earnings surprise and negatively related to the earnings loss dummy, etc., in this specification. Looking at positive and negative coverage in this simple way, without reference to any of the firm-specific or market-wide determinants that are significant in predicting coverage, results in a trading strategy that is equivalent to earnings momentum (albeit one with relatively high risk-adjusted returns).

As we proceed to include additional cross-sectional predictors of coverage in Panels B, C, D, E, and F, an interesting pattern emerges. Evidence of underreaction to positive earnings news generally persists, but it wanes with respect to the high PMC^+ stocks (high coverage good news) while staying robust for the low PMC^- firms (low coverage good news). In other words, as we move away from a simpler specification of coverage, the evidence of underreaction to high-coverage positive events disappears while the underreaction with respect to low-coverage positive events persists.

The results strongly suggest that a risk-based factor such as size or liquidity, which we would expect to impact both low PMC^+ and low PMC^- firms in similar ways, cannot explain this asymmetry between low-coverage and high-coverage events. In particular, the lowest decile of PMC^+ stocks never have alphas significantly different from zero, as we would expect to see if all low-coverage firms were being rewarded for loading on some symmetric risk factor (such as information risk. If, instead, one were to argue that high expected returns for low-coverage bad news stocks were being offset by earnings momentum, resulting in no net predictability for the low PMC$^+$ stocks, then we would surely expect to see evidence of such earnings-related drift in panel A, where PMC1$^+$ and PMC1$^-$ only load on positive and negative earnings news attributes. On the contrary, in each panel, we only find evidence of underreaction with respect to positive events, an effect which becomes increasingly concentrated on low-coverage events as we include additional firm-specific predictors of coverage behavior.

At the same time, as we move toward the more complete specifications of predicted coverage, the predictability of absolute PMC begins to look more like a weaker version of PMC$^-$, suggesting that the underreaction to low-coverage positive events exhibited by PMC$^-$ is behind the observed predictability of absolute PMC.

Panels G and H show estimated month t+n excess returns for portfolios formed on PMC5 and PMC6, respectively. In particular, Panel H illustrates that the asymmetric predictability results are somewhat persistent over subsequent months. For example, estimated four-factor alphas for the long-short PMC6$^-$ portfolios fall relatively sharply after month t+1, but remain positive and statistically significant at 5 percent out to month t+5, addressing the potential concern that we might simply be observing return predictability based on residual short-term earnings drift stemming from announcements made towards the ends of the 3-month formation periods.

8 Conclusion

Utilizing estimates of financial news coverage as a proxy for investor attention, this study identifies asymmetric delays in the market's response to positive news events as a potential explanation for the neglected firm effect. I identify the cross-sectional event- and firm-specific characteristics that predict positive and negative media coverage regarding almost 180,000 quarterly earnings announcements from 1984 to 2005, showing that positive corporate news tends to go relatively unnoticed compared to negative news regarding otherwise-similar firms. If news media outlets maximize readership by attempting to publish stories that readers will find most interesting, the results suggest that the impact of cognitive constraints (i.e., limited attention) on investors' information preferences is asymmetric with respect to positive and negative news events. If limits to arbitrage are binding, this relative lack of attention with respect to some events will contribute to predictability in asset returns. Finally, if good news from neglected firms is more likely to be ignored, low media coverage will predict positive returns for these stocks, on average. In short, my findings support this hypothesis.

This study contributes to the existing literature in the following ways: First, examining the cross-sectional news story elements that predict coverage in the media, I find evidence of significant asymmetry in reporting: negative earnings information is more likely to result in media coverage than is positive information, holding other factors constant; second, I propose and apply a new measure of investor attention based on the predictable component of media coverage decisions: probability of media coverage (PMC); third, I utilize PMC to identify relatively "neglected" stocks, confirming the results of other studies which find that these firms enjoy a return premium that cannot be explained by the standard set of risk factors. In contrast to these earlier papers, however, I find that this apparent neglected firm effect is attributable to systematic underreaction to positive news for these firms (i.e., consistent with the effects of negativity bias in attention), while the stocks of relatively neglected firms with negative news appear to be efficiently priced.

Regarding this evidence of asymmetry in attention, why might the media (and, by implication, investors) be more likely to focus on negative events rather than positive ones? If the market for news is reasonably competitive, then this observed bias in coverage must, in some sense, reflect the underlying information preferences/requirements of financial news readers. While some potential explanations for negativity bias are identified in the literature, the answer to this question is beyond the scope of the current paper; this remains a potentially important line of inquiry for future research.

While the study focuses on an examination of media coverage with respect to one particular type of news event (albeit one that lends itself particularly well to the question of interest), the results underscore the potential importance of investor attention allocations for our understanding of the market's reaction to news.

References

Amihud, Yakov, and Haim Mendelson, 1986, Asset Pricing and the Bid-Ask Spread, *Journal of Financial Economics* 17(2), 223-249.

Antweiler, Werner, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *Journal of Finance* 59, 1259-1294.

Antweiler, Werner, and Murray Z. Frank, 2006, Do U.S. Stock Markets Typically Overreact to Corporate News Stories, Working paper, University of British Columbia and University of Minnesota.

Arbel, Avner, and Paul Strebel, 1982, The Neglected and Small Firm Effects, *Financial Review* 17, 201-218.

Ball, Ray, and Philip Brown, 1968, An Empirical Evaluation of Accounting Income Numbers, *Journal of Accounting Research* 6 (2), 159-178.

Bamber, Linda Smith, 1987, Unexpected Earnings, Firm Size, and Trading Volume Around Quarterly Earnings Announcements, *Accounting Review* 62 (3), 510-532.

Banz, R.W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3–18.

Barber, Brad M., and Terrance Odean, 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.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.

Basak, Suleyman, and Domenico Cuoco, 1998, An Equilibrium Model with Restricted Stock Market Participation, *Review of Financial Studies* 11 (2), 309-341.

Bernard, Victor L., and Jacob K. Thomas, 1990, Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium, *Journal of Accounting Research* 27, 1-36.

Bhattacharya , Utpal, Neal Galpin, Rina Ray, and Xiaoyun Yu, 2006, The Role of the Media in the Internet IPO Bubble, *Journal of Financial and Quantitative Analysis*, forthcoming.

Brennan, Michael J., Narasimhan Jegadeesh, and Bhaskaran Swaminathan, 1993, Investment Analysis and the Adjustment of Stock Prices to Common Information, *Review of Financial Studies* 6(4), 799-824.

Busse, Jeffrey A., and T. Clifton Green, 2002, Market efficiency in real time, *Journal of Financial Economics* 65, 415-437.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Chan, Wesley S., 2003, Stock price reaction to news and no-news: drift and reversal after headlines, *Journal of Financial Economics* 70, 223-260.

Christensen, Theodore E., Toni Q. Smith, and Pamela S. Stuerke, 2004, Public Predisclosure Information, Firm Size, Analyst Following, and Market Reactions to Earnings Announcements, *Journal of Business Finance & Accounting* 31 (7-8), 951-984.

Conrad, Jennifer, Bradford Cornell, and Wayne R. Landsman, 2002, When Is Bad News Really Bad News?, *Journal of Finance* 57 (6), 2507-2532.

Conrad, Jennifer, Bradford Cornell, Wayne R. Landsman, and Brian Rountree, 2004, How do Analyst Recommendation Respond to Major News?, Working paper, University of North Carolina – Chapel Hill, University of California – Los Angeles, and Rice University.

DellaVigna, Stefano, and Joshua Pollet, 2004, Strategic release of information on Friday: Evidence from earnings announcements, Working paper, University of California at Berkeley and University of Illinois at Urbana-Champaign.

DeLong, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379-395.

Dyck, Alexander, and Luigi Zingales, 2003, The media and asset prices, Working paper, Harvard Business School and University of Chicago.

Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57(5), 2185-2221.

Engelberg, Joseph, 2007, Costly Information Processing: Evidence from Earnings Announcements, Working paper, Kellogg School of Management, Northwestern University.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fang, Lily, and Joel Peress, 2007, Media Coverage and the Cross-Section of Returns, Working paper, INSEAD.

Foerster, Stephen R., and G. Andrew Karolyi, 1999, The Effects of Market Segmentation and Investor Recognition on Asset Prices: Evidence from Foreign Stocks Listing in the United States, *Journal of Finance* 54 (3), 981-1013.

Gadarowski, Christopher, 2002, Financial press coverage and expected stock returns, Working paper.

Harrington, David E., 1989, Economic news on television: The determinants of coverage, *Public Opinion Quarterly* 53, 17-40.

Hirshleifer, David, Sonya S. Lim, and Siew H. Teoh, 2006, Driven to distraction: Extraneous events and underreaction to earnings news, Working paper, University of California at Irvine, and Depaul University.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55 (1), 265-295.

Hou, Kewei, and Tobias J . Moskowitz, 2005, Market Frictions, Price Delay, and the Cross-Section of Expected Returns, *Review of Financial Studies* 18 (3), 981-1020.

Hou, Kewei, Lin Peng, and Wei Xiong, 2006, A tale of two anomalies: The implications of investor attention for price and earnings momentum, Working paper, Ohio State University, Baruch College, and Princeton University.

Huang, Lixin, and Hong Liu, 2007, Rational Inattention and Portfolio Selection, *Journal of Finance* 62 (4), 1999-2040.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer: A nonevent that made stock prices soar, *Journal of Finance* 56, 387-396.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk, *Econometrica* 47, 263-291.

Kaniel, Ron, Laura Starks, and Vasudha Vasudevan, 2007, Headlines and Bottom Lines: Attention and Learning from Media Coverage of Mutual Funds, Working paper, Duke University, and University of Texas at Austin.

Kothari, S. P., Susan Shu, Peter Wysocki, 2005, Do Managers Withhold Bad News?, Working paper, MIT Sloan and Boston College.

Manning, Christopher D., and Hinrich Schuetze, 1999, Foundations of Statistical Natural Language Processing, MIT Press, Cambridge, MA.

McCallum, Andrew Kachites, 1996, *Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering*. http://www.cs.cmu.edu/ mccallum/bow, 1996.

McQueen, Grant, Michael Pinegar, and Steven Thorley, 1996, Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns, *Journal of Finance* 51 (3), 889-919.

Merton, Robert C., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42 (3), 483-510.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111(3), 642-685.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence, and categorical learning, *Journal of Financial Economics* 80(3), 563-602.

Reinganum, M. R., 1981, Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earning Yields and Market Values, *Journal of Financial Economics* 9, 19-46.

Sadka, Ronnie, 2006, Momentum and post-earnings announcement drift anomalies: The role of liquidity risk, *Journal of Financial Economics* 80(2), 309-349.

Shapiro, Alexander, 2002, The Investor Recognition Hypothesis in a Dynamic General Equilibrium: Theory and Evidence, *Review of Financial Studies* 15 (1), 97-141.

Skinner, Douglas J., and Richard G. Sloan, 2002, Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio, *Review of Accounting Studies* 7, 289-312.

Soroka, Stuart N., 2006, Good news and bad news: Asymmetric responses to economic information, *Journal of Politics* 68 (2), 372-385.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of the media in the stock market, *Journal of Finance* 62 (3), 1139-1168.

Tetlock, Paul C., Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63 (3), 1437-1467.

Veronesi, Pietro, 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, *Review of Financial Studies* 12 (5), 975-1007.

Appendix A: Classification of news articles

In building the media data set, 68,102 Wall Street Journal news articles with Intelligent Indexing Code c151 ("earnings") were obtained from Factiva in text format from October 1984 to December 2005.[13] Articles were then classified using the computational linguistics program Rainbow (McCallum (1996)), applying a two-stage, Naïve Bayesian, bag-of-words methodology. In order to "train" the classification algorithm, 500 articles were selected at random for hand-classification. Of these, 129 were identified as "not earnings articles" (i.e., articles that did not discuss the content of a specific recent corporate earnings announcement), and 371 were identified as "earnings articles" (stage 1). Further categorizing the set of "earnings articles", 147 were classified as "negative", and 224 as "not negative" (stage 2). The classification software was then applied to the articles, resulting in predictive models of category membership based on word frequencies. The first binary classification model (WSJ1) distinguishes between "earnings" and "not earnings" articles using odds-ratios from the top 300 bigrams ranked by infogain. The second binary classification model (WSJ2) distinguishes between "negative" and "not negative" among identified "earnings" articles using odds-ratios from the top 25 unigrams ranked by infogain.

Model accuracy was tested by randomly excluding 100 articles from the training set, re-estimating the model, and then examining accuracy with respect to the classifications predicted for the excluded articles. In the course of 100 such trials, model WSJ1 had an average accuracy of 88.52%, while model WSJ2 implied an average accuracy of 83.42%. This estimated rate of accuracy is comparable to hand-classification methods using paid assistants.

Models WSJ1 and WSJ2 were then applied to the full data set of uncategorized articles to calculate estimated probabilities of category membership for each. Articles with predicted probability >0.5 of being an "earnings" article were recorded as such. Of these, articles with a predicted probability >0.5 from WSJ2 of being a "negative" article were identified as "negative earnings", with all others recorded as "not negative earnings". The categorization results are shown in Table B1.

[13] Additional filters were also applied in excluding the most common types of "false positive" articles resulting from the Factiva search – details are available upon request.

Table A1: WSJ full and training sample article classification

	"Not Earnings"	"Earnings"	
Training Sample – Classified by hand n=500	129 (25.8%)	371 (74.2%)	
		"Not Negative"	"Negative"
		224 (60.38%)	147 (39.6%)

	"Not Earnings"	"Earnings"	
Full Sample – Classified by algorithm (WSJ1 and WSJ2) n=68,102	19,989 (29.352%)	49,113 (70.648%)	
		"Not Negative"	"Negative"
		31,833 (64.816%)	17,280 (35.184%)

<u>Example of a WSJ article classified as "Positive earnings announcement news":</u>

HD Walgreen Co. Earnings Rise 17% on Sales Gain Of Prescription Drugs
WC 194 words
PD 4 January 2002
LP DEERFIELD, Ill. -- Walgreen Co., boosted by strong prescription sales, said fiscal first-quarter earnings jumped 17%.
The drugstore chain reported net income of $185.9 million, or 18 cents a share, compared with $158.4 million, or 15 cents a share, a year earlier. Analysts surveyed by Thomson Financial/First Call had forecast earnings of 17 cents a share for the period ended Nov. 30.
TD The recent quarter's results included a $5.5 million pretax gain from the final payment of an antitrust settlement regarding brand-name prescription drugs. Walgreen said prescriptions, which accounted for 60% of first-quarter sales, leapt 22% overall and 17% on a same-store basis. Total sales for the period climbed 17% to $6.6 billion from $5.6 billion. Chairman L. Daniel Jorndt said the chain planned to open 475 new stores and two new distribution centers this year. As of Nov. 30, the company operated 3,623 drugstores in 43 states and Puerto Rico. Shares of Walgreen rose $1.33 to $34.41 as of 4 p.m. in New York Stock Exchange composite trading.

<u>Example of a WSJ article classified as "Negative earnings announcement news":</u>

HD Business Brief -- Playtex Products Inc.: Loss of $959,000 Is Posted Amid a 7.9% Decrease in Sales
WC 185 words
PD 10 February 2004
LP Playtex Products Inc. posted a fourth-quarter loss as the consumer-products company was hurt by restructuring charges, competition in the tampon market and weather that cut into sales of sun-care products. The Westport, Conn., company reported a fourth-quarter loss of $959,000, or two cents a share, including restructuring charges of five cents. A year earlier, the company reported a profit of $7.2 million, or 12 cents a share. Playtex, which makes household items such as Playtex tampons, Wet Ones wipes and Banana Boat sunscreen, said fourth-quarter sales fell 7.9% to $146.7 million from $159.2 million a year earlier. Playtex, which faces intense competition from Procter & Gamble Co., had lowered its 2003 and 2004 outlook on Jan. 30, projecting a fourth-quarter net loss of two cents a share. The results were issued after the close of regular trading. In 4 p.m. composite trading yesterday on the New York Stock Exchange, Playtex was unchanged at $6.30.

Figure 1: Firm and event characteristics that attract (absolute) media coverage – Illustration based on the results from Table 2

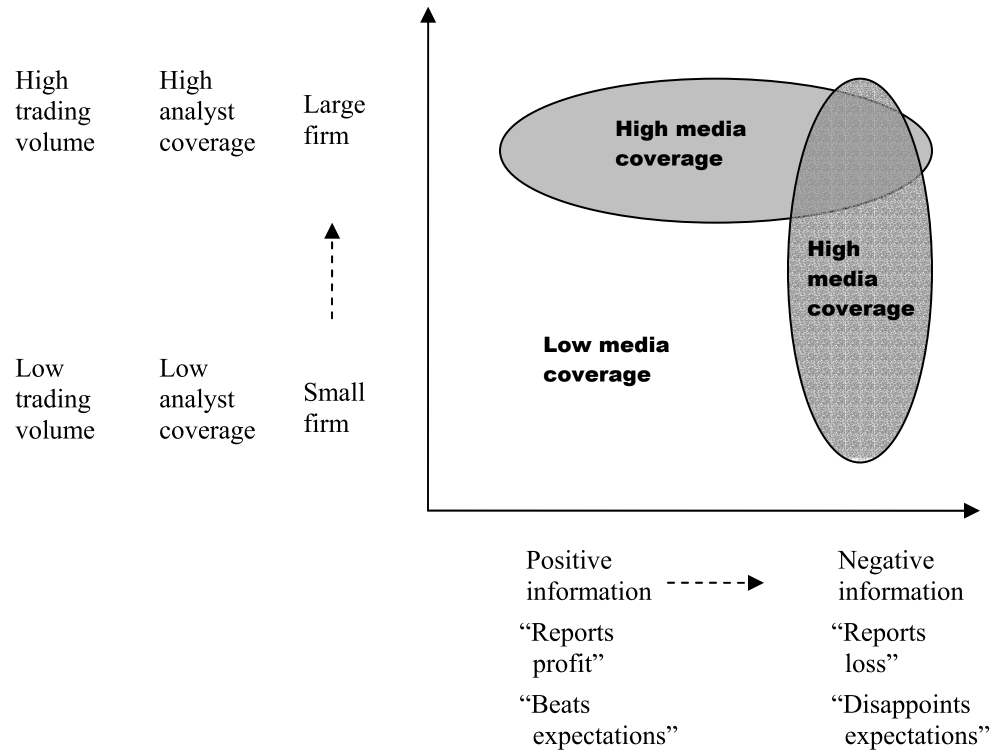

Figure 2: Empirical distributions of probability of media coverage (PMC)

(200 bins); Probability of coverage (PMC) is the sum of the probabilities of positive coverage (PMC^{+}) and negative coverage (PMC^{-})

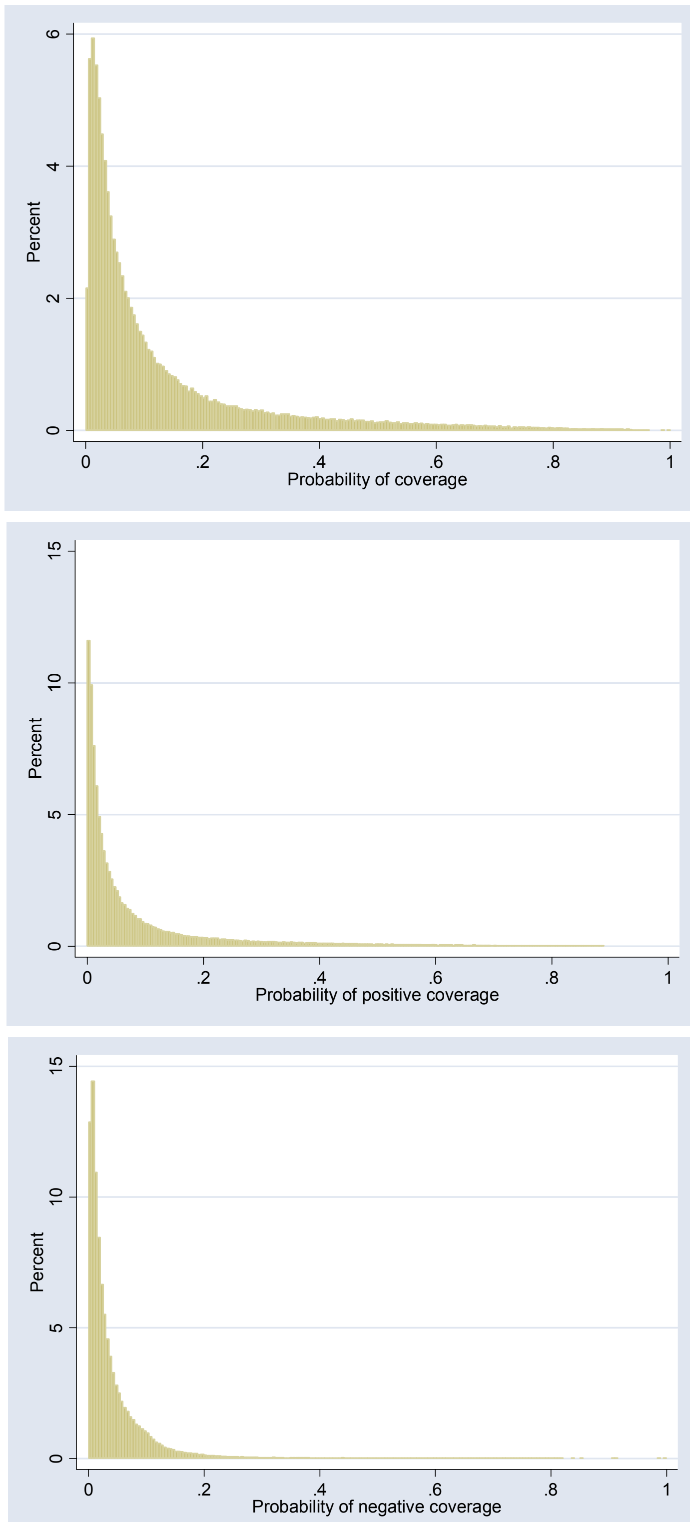

Figure 3: Probability of media coverage (PMC) by Fama-French 49 industry classification

This figure shows the bottom-15 and top-15 Fama-French 49 industries ranked by average PMC.

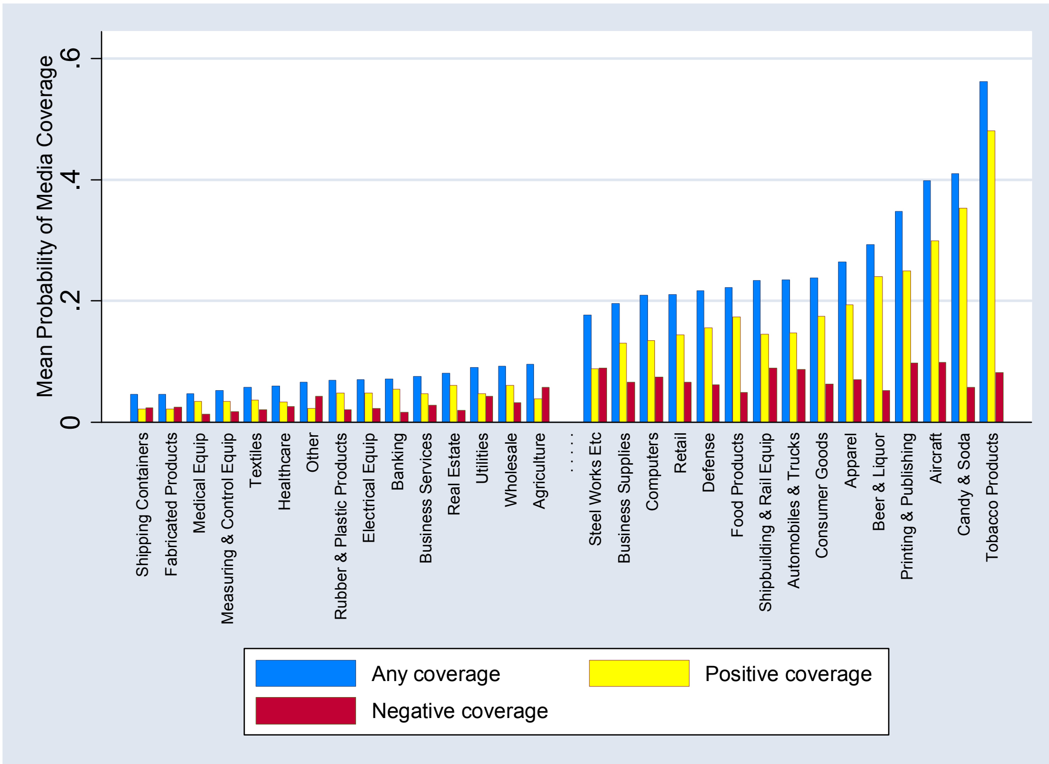

Table 1: Descriptive statistics of earnings announcement information variables and firm characteristics by observed media coverage

This table reports descriptive statistics for earnings announcements in the sample. *Earnings surprise* is the announced EPS less the median analyst forecast from the 30 days prior to the earnings announcement, divided by the stock price. *Analyst attention* is the number of distinct analyst EPS forecasts observed in the 30 days prior to the announcement. *stdev(forecasts)* is the standard deviation of analysts' EPS forecasts (normalized by the stock price) observed in the 30 days prior to the announcement. *Market value* is the number of shares outstanding multiplied by the closing stock price two days before the announcement. *$Trading Volume* is the firm's average dollar value of trading volume in the 60 trading days prior to the announcement. *B/M* is the book value of equity from the fiscal year ending in the previous year divided by the market value of equity from December 31 (divided by 1000). *US firm* is a dummy variable that is equal to 1 if the firm is identified as a US firm in CRSP. *Beta* is the estimated slope coefficient from a CAPM regression of daily returns from the 60 days prior to the announcement window. *Institutional ownership* is the percentage of equity held by institutions at the end of the previous calendar year. *Recent returns* is the average daily stock return during the 60 trading days prior to the announcement window. *stdev(Recent returns)* is the standard deviation of daily stock returns during the 60 trading days prior to the announcement window. *CAR* is the abnormal announcement-window stock return based on a four-factor model of daily returns estimated over the previous 60 trading days. Panel A reports statistics for all earnings announcements in the sample. Panel B reports statistics only for those announcements with an associated positive Wall Street Journal news article. Panel C reports statistics for those announcements without any associated Wall Street Journal news article. Panel D reports statistics only for those announcements with an associated negative Wall Street Journal news article.

Panel A: All announcements

variable	*N*	*mean*	*sd*	*min*	*p1*	*p5*	*p10*	*p25*	*p50*	*p75*	*p90*	*p95*	*p99*	*max*
Earnings surprise	190808	-0.00252	0.038373	-0.99429	-0.09143	-0.01474	-0.00571	-0.00078	0	0.001096	0.00354	0.006923	0.030303	0.977966
Earnings (%)	190461	0.005336	0.120014	-0.99921	-0.19033	-0.03688	-0.01031	0.002971	0.0088	0.015962	0.024583	0.032537	0.071048	25.04173
Analyst attention	190808	6.419044	5.112228	2	2	2	2	3	5	8	13	17	25	46
stdev(forecasts)	190808	0.210451	0.772615	0	0	0	0	0.02439	0.0625	0.166667	0.421053	0.833333	2.5	95.60001
Market value	190808	3124359	1.34E+07	1148.624	23174.99	50998.24	80076.49	186213.8	548087.6	1765107	5571076	1.15E+07	4.79E+07	5.79E+08
$Trading Volume	190807	1.60E+07	7.28E+07	63.11498	33728.63	109670.9	204817.3	617034.9	2339028	9021924	2.99E+07	6.15E+07	2.33E+08	4.42E+09
B/M	181249	0.016978	1.195242	3.40E-08	2.08E-05	7.38E-05	0.00012	0.000236	0.000416	0.000663	0.000977	0.001299	0.002806	186.9693
US firm	190808	0.97564	0.154163	0	0	1	1	1	1	1	1	1	1	1
Beta	190808	0.806831	0.726316	-5.0086	-0.73189	-0.1568	0.047852	0.351605	0.721025	1.16065	1.696	2.1108	3.0715	7.9255
Institutional ownership (%)	183749	46.63316	23.00282	0.000172	2.747857	9.564054	15.299	28.31198	46.83545	64.41245	77.43977	84.15045	93.88927	99.99456
Recent returns [-61 to -2]	190808	0.071631	0.407534	-3.67218	-1.09541	-0.5746	-0.36752	-0.11962	0.075913	0.269385	0.498527	0.688497	1.195283	5.973645
stdev(Recent returns)	190808	0.028847	0.017571	0	0.007895	0.010625	0.012553	0.016827	0.024312	0.035775	0.050508	0.062259	0.091454	0.35137
CAR[-1 to 3]	190808	0.091849	9.55053	-147.174	-27.8574	-14.4875	-9.67611	-3.98083	0.070352	4.3158	9.90672	14.45842	26.7681	218.69

Panel B: Announcements with positive coverage in the WSJ

variable	*N*	*mean*	*sd*	*min*	*p1*	*p5*	*p10*	*p25*	*p50*	*p75*	*p90*	*p95*	*p99*	*max*
Earnings surprise	16901	0.000234	0.010242	-0.46	-0.01226	-0.0034	-0.00134	0	0.000193	0.000879	0.002505	0.00451	0.013333	0.228571
Earnings (%)	16899	0.013149	0.117317	-0.36429	-0.00491	0.000925	0.001882	0.00456	0.009412	0.016053	0.024776	0.032363	0.061359	15
Analyst attention	16901	11.75481	6.63597	2	2	3	4	7	11	16	21	24	30	44
stdev(forecasts)	16901	0.108481	0.324927	0	0	0	0	0.020408	0.045455	0.1	0.2	0.333333	1.05	13.66667
Market value	16901	1.34E+07	3.00E+07	13793.5	93785.72	277312	527963.1	1494986	4446590	1.23E+07	3.02E+07	5.46E+07	1.50E+08	5.26E+08
$Trading Volume	16901	5.93E+07	1.41E+08	6425.677	231694.7	965674.6	1925451	6012078	1.90E+07	5.54E+07	1.40E+08	2.29E+08	6.31E+08	3.03E+09
B/M	16496	0.047821	2.458221	2.20E-07	2.72E-05	7.02E-05	0.000109	0.000199	0.000347	0.00055	0.000792	0.000987	0.001666	186.9693
US firm	16901	0.975564	0.154404	0	0	1	1	1	1	1	1	1	1	1
Beta	16901	0.935877	0.584175	-2.5804	-0.21561	0.1321	0.29873	0.56961	0.87782	1.2231	1.6167	1.9475	2.83	5.0124
Institutional ownership (%)	16545	56.3496	19.00853	0.049414	6.463413	21.03038	29.7944	44.86758	58.0683	70.10284	79.30531	84.69053	93.54565	99.98873
Recent returns [-61 to -2]	16901	0.08432	0.311557	-1.90529	-0.78797	-0.40647	-0.25956	-0.07294	0.084333	0.243917	0.418465	0.560178	0.94692	3.293368
stdev(Recent returns)	16901	0.02208	0.011839	0.001478	0.007914	0.010008	0.011457	0.014327	0.018954	0.026534	0.036365	0.044091	0.064262	0.32312
CAR[-1 to 3]	16901	0.659712	7.620515	-85.7621	-21.5445	-10.5094	-6.8927	-2.88749	0.4863	4.16483	8.86934	12.43585	22.3734	93.3546

Panel C: Announcements with no coverage in the WSJ

variable	*N*	*mean*	*sd*	*min*	*p1*	*p5*	*p10*	*p25*	*p50*	*p75*	*p90*	*p95*	*p99*	*max*
Earnings surprise	165732	-0.0024	0.038135	-0.99429	-0.08833	-0.015	-0.00585	-0.00082	0	0.001131	0.003636	0.007111	0.031189	0.977966
Earnings (%)	165423	0.0054	0.121556	-0.99921	-0.19	-0.03833	-0.01143	0.002977	0.008889	0.016058	0.024675	0.032668	0.072533	25.04173
Analyst attention	165732	5.707262	4.429997	2	2	2	2	3	4	7	11	15	22	46
stdev(forecasts)	165732	0.210688	0.781761	0	0	0	0	0.02439	0.0625	0.166667	0.421053	0.833333	2.5	95.60001
Market value	165732	1922393	9292109	1148.624	21861.39	47784	73454.42	163409.8	446050.3	1268752	3417051	6335140	2.35E+07	5.79E+08
$Trading Volume	165731	1.05E+07	5.78E+07	63.11498	30981.78	98841.34	182152.2	522420	1833569	6394109	1.91E+07	3.68E+07	1.41E+08	4.42E+09
B/M	156819	0.012222	0.861688	3.40E-08	2.06E-05	7.37E-05	0.000121	0.000238	0.000418	0.000664	0.000975	0.00129	0.002763	186.9693
US firm	165732	0.975756	0.153806	0	0	1	1	1	1	1	1	1	1	1
Beta	165732	0.78631	0.73744	-5.0086	-0.77203	-0.18597	0.024088	0.32439	0.69311	1.1436	1.7013	2.1208	3.0785	7.9255
Institutional ownership (%)	159246	45.30571	23.20005	0.000172	2.546038	8.973255	14.41449	26.68684	44.69246	63.15168	77.07803	84.12677	94.01344	99.99456
Recent returns [-61 to -2]	165732	0.073242	0.412661	-3.67218	-1.10158	-0.58314	-0.37369	-0.12196	0.076681	0.274004	0.508065	0.701447	1.216335	5.973645
stdev(Recent returns)	165732	0.029473	0.0178	0	0.00786	0.010682	0.012705	0.017229	0.025031	0.036667	0.051418	0.063262	0.092352	0.35137
CAR[-1 to 3]	165732	0.102713	9.678056	-147.174	-27.9506	-14.6799	-9.84426	-4.05451	0.05957	4.382648	10.09807	14.75812	27.25886	218.69

Panel D: Announcements with negative coverage in the WSJ

variable	*N*	*mean*	*sd*	*min*	*p1*	*p5*	*p10*	*p25*	*p50*	*p75*	*p90*	*p95*	*p99*	*max*
Earnings surprise	8175	-0.01079	0.067714	-0.98824	-0.31563	-0.05807	-0.02177	-0.0041	0	0.000898	0.00417	0.009534	0.055118	0.82
Earnings (%)	8139	-0.01218	0.087904	-0.95915	-0.44353	-0.11333	-0.04704	-0.00604	0.004378	0.012851	0.021767	0.029818	0.065494	2.289157
Analyst attention	8175	9.817859	6.358488	2	2	2	3	5	9	13	19	22	29	41
stdev(forecasts)	8175	0.41645	1.122849	0	0	0	0.021277	0.055556	0.142857	0.363636	1	1.5	4.5	33
Market value	8175	6327817	1.80E+07	2697.2	34312	94783.49	174495.8	557669.8	1808531	5281947	1.41E+07	2.46E+07	7.56E+07	4.35E+08
$Trading Volume	8175	3.67E+07	1.01E+08	5005	77214.07	327755.3	710128.9	2626058	9466614	3.05E+07	8.13E+07	1.56E+08	4.66E+08	2.32E+09
B/M	7934	0.046857	2.322482	2.08E-07	2.24E-05	9.52E-05	0.00016	0.000314	0.000553	0.000877	0.001422	0.002006	0.006169	141.275
US firm	8175	0.973456	0.160757	0	0	1	1	1	1	1	1	1	1	1
Beta	8175	0.95605	0.719325	-3.6665	-0.54913	-0.01279	0.20408	0.51837	0.87327	1.2906	1.8095	2.3072	3.2245	6.0908
Institutional ownership (%)	7958	52.99563	20.58458	0.149269	5.001407	15.36306	23.01242	39.11233	55.2866	68.79938	77.8599	83.23251	92.13445	99.49509
Recent returns [-61 to -2]	8175	0.012745	0.468732	-3.54245	-1.46057	-0.73001	-0.49402	-0.19432	0.035713	0.239437	0.47199	0.672202	1.242535	3.77118
stdev(Recent returns)	8175	0.030143	0.019721	0.002262	0.008574	0.011433	0.013281	0.017254	0.024349	0.036978	0.054154	0.067915	0.10369	0.27084
CAR[-1 to 3]	8175	-1.30241	10.36446	-115.566	-33.2927	-17.8895	-11.709	-5.2624	-0.7014	3.42588	8.63563	12.982	25.1824	146.618

Table 2: Which factors predict media coverage?

This table reports coefficient estimates from logit regressions of *ABS(COVERAGE)*$_{i,t}$ on market conditions and firm and event characteristics; *ABS(COVERAGE)*$_{i,t}$ is equal to 1 if there is an identified positive or negative WSJ earnings article within one week of the announcement date, and 0 otherwise; *Earnings surprise* is the announced EPS less the median analyst forecast from the 30 days prior to the earnings announcement divided by the stock price; *I{Negative surprise}* is a dummy variable equal to one if *Earnings surprise* is negative, and zero otherwise; *I{Loss}* is a dummy variable equal to one if announced EPS is negative, and zero otherwise; *log(Market value)* is the natural logarithm of the number of shares outstanding multiplied by the closing stock price two days before the announcement; *B/M* is the book value of equity from the fiscal year ending in the previous year divided by the market value of equity from December 31 (divided by 1000); *Institutional ownership* is the percentage of equity held by institutions at the end of the previous calendar year; *log*(*$Trading Volume)* is the natural logarithm of the firm's average dollar value of trading volume in the 60 trading days prior to the announcement; *log(Analyst attention)* is the natural log of the number of distinct analyst forecasts observed in I/B/E/S in the 30 days preceding the announcement; *stdev(forecasts)* is the standard deviation of the normalized analyst EPS forecasts recorded in I/B/E/S in the 30 days preceding the announcement; *Recent returns* is the average daily stock return during the 60 trading days prior to the announcement window; *S&P500 returns* is the average daily return on the S&P500 during the 60 trading days prior to the announcement window; *stdev(S&P500 returns)* is the standard deviation of the average daily return on the S&P500 during the 60 trading days prior to the announcement window; month-of-the-year and day-of-the-week dummy variables are included in all specifications; significant standard errors in brackets, robust to arbitrary heteroskedasticity and clustering by firm; * significant at 10%; ** significant at 5%; *** significant at 1%.

	Media coverage dummy			
	(1)	(2)	(3)	(4)
Earnings surprise	-1.918***	-1.869***	-1.954***	-1.915***
	[0.256]	[0.267]	[0.271]	[0.282]
I{Negative surprise}	0.118***	0.062**	0.122***	0.074***
	[0.026]	[0.026]	[0.026]	[0.025]
I{Loss}	0.241***	0.358***	0.392***	0.509***
	[0.066]	[0.068]	[0.067]	[0.069]
I{Negative surprise}*I{Loss}	0.242***	0.202***	0.197***	0.158**
	[0.064]	[0.065]	[0.065]	[0.067]
log(Market value)	0.524***	0.470***	0.543***	0.489***
	[0.031]	[0.031]	[0.034]	[0.035]
B/M (/1000)	0.013***	0.014***	0.01	0.012*
	[0.003]	[0.003]	[0.008]	[0.007]
Institutional ownership (%)	0.002	0.004***	-0.001	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
log(Value of trading volume [-61 to -2])	0.090***	0.203***	0.121***	0.234***
	[0.026]	[0.027]	[0.027]	[0.029]
log(Analyst attention)	0.513***	0.457***	0.497***	0.445***
	[0.044]	[0.044]	[0.044]	[0.044]
stdev(Analyst forecasts)	0.077***	0.067***	0.073***	0.065***
	[0.016]	[0.015]	[0.016]	[0.016]
Recent returns [-61 to -2]	-0.005***	-0.005***	-0.005***	-0.005***
	[0.000]	[0.000]	[0.000]	[0.000]
S&P500 returns [-61 to -2]	0.016***	0.006***	0.016***	0.005***
	[0.001]	[0.001]	[0.001]	[0.001]
stdev(S&P500 returns [-61 to -2])	-0.04	0.053*	-0.029	0.054*
	[0.028]	[0.028]	[0.029]	[0.029]
Year dummies	No	Yes	No	Yes
Industry dummies	No	No	Yes	Yes
Observations	178898	178898	178898	178898
Pseudo R-squared	0.21	0.22	0.24	0.25

Table 3: Determinants of positive and negative media coverage

This table reports coefficient estimates from multinomial logit regressions of $COVERAGE_{i,t}$ on market conditions and firm and event characteristics. $COVERAGE_{i,t}$ is equal to 1 if there is an identified positive WSJ earnings article within one week of the announcement date, -1 if there is an identified negative WSJ earnings article within one week of the announcement date, and 0 otherwise; *Earnings surprise* is the announced EPS less the median analyst forecast from the 30 days prior to the earnings announcement divided by the stock price; *I{Negative surprise}* is a dummy variable equal to one if *Earnings surprise* is negative, and zero otherwise; *I{Loss}* is a dummy variable equal to one if announced EPS is negative, and zero otherwise; *log(Market value)* is the natural logarithm of the number of shares outstanding multiplied by the closing stock price two days before the announcement; *B/M* is the book value of equity from the fiscal year ending in the previous year divided by the market value of equity from December 31 (divided by 1000); *Institutional ownership* is the percentage of equity held by institutions at the end of the previous calendar year; *log($Trading Volume)* is the natural logarithm of the firm's average dollar value of trading volume in the 60 trading days prior to the announcement; *log(Analyst attention)* is the natural log of the number of distinct analyst forecasts observed in I/B/E/S in the 30 days preceding the announcement; *stdev(forecasts)* is the standard deviation of the normalized analyst EPS forecasts recorded in I/B/E/S in the 30 days preceding the announcement; *Recent returns* is the average daily stock return during the 60 trading days prior to the announcement window; *S&P500 returns* is the average daily return on the S&P500 during the 60 trading days prior to the announcement window; *stdev(S&P500 returns)* is the standard deviation of average daily return on the S&P500 during the 60 trading days prior to the announcement window; results for year, month-of-the-year, day-of-the-week and industry dummies not shown; standard errors in brackets, robust to arbitrary heteroskedasticity and clustering by firm; * significant at 10%; ** significant at 5%;*** significant at 1%

	Negative coverage	*Positive coverage*	*Negative coverage*	*Positive coverage*	*Negative coverage*	*Positive coverage*	*Negative coverage*	*Positive coverage*	*Negative coverage*	*Positive coverage*
	(1)		(2)		(3)		(4)		(5)	
Earnings surprise	-1.958***	0.682	-1.906***	0.806	-1.874***	0.911	-1.838***	1.102	-1.859***	0.973
	[0.243]	[0.958]	[0.246]	[0.982]	[0.254]	[0.953]	[0.256]	[0.968]	[0.255]	[0.932]
I{Negative surprise}	0.701***	-0.065**	0.630***	-0.123***	0.672***	-0.047	0.612***	-0.095***	0.616***	-0.101***
	[0.035]	[0.030]	[0.035]	[0.030]	[0.036]	[0.030]	[0.036]	[0.029]	[0.036]	[0.029]
I{Loss}	1.574***	-1.674***	1.645***	-1.497***	1.726***	-1.534***	1.785***	-1.359***	1.772***	-1.382***
	[0.070]	[0.134]	[0.072]	[0.134]	[0.070]	[0.134]	[0.072]	[0.134]	[0.069]	[0.133]
I{Negative surprise}*I{Loss}	-0.356***	0.486***	-0.367***	0.421***	-0.403***	0.453***	-0.410***	0.392***	-0.394***	0.425***
	[0.067]	[0.147]	[0.068]	[0.148]	[0.068]	[0.148]	[0.069]	[0.149]	[0.067]	[0.145]
log(Market value of equity)	0.407***	0.581***	0.339***	0.460***	0.412***	0.603***	0.347***	0.483***	0.335***	0.473***
	[0.037]	[0.043]	[0.038]	[0.044]	[0.038]	[0.043]	[0.039]	[0.044]	[0.037]	[0.042]
B/M (/1000)	0.017*	0.015***	0.019**	0.016***	0.015	0.012	0.017	0.013		
	[0.009]	[0.004]	[0.008]	[0.005]	[0.013]	[0.009]	[0.011]	[0.009]		
Institutional ownership (%)	0.001	0.001	0.003**	0.004***	-0.001	-0.002	0.001	0		
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]		
log(Value of trading volume [-61 to -2])	0.108***	0.119***	0.227***	0.301***	0.159***	0.146***	0.276***	0.329***	0.283***	0.338***
	[0.031]	[0.035]	[0.032]	[0.038]	[0.031]	[0.035]	[0.033]	[0.038]	[0.030]	[0.035]
US firm dummy	0.134	-0.034	-0.094	-0.353**	0.166	-0.043	-0.066	-0.373**	-0.066	-0.351**
	[0.203]	[0.164]	[0.198]	[0.159]	[0.192]	[0.168]	[0.193]	[0.174]	[0.180]	[0.162]
log(Analyst attention)	0.568***	0.465***	0.495***	0.404***	0.542***	0.449***	0.469***	0.392***	0.484***	0.383***

	[0.049]	[0.049]	[0.050]	[0.050]	[0.049]	[0.050]	[0.049]	[0.050]	[0.047]	[0.048]
stdev(Analyst forecasts)	0.109***	-0.009	0.099***	-0.023	0.102***	-0.001	0.094***	-0.004	0.092***	-0.004
	[0.019]	[0.028]	[0.019]	[0.029]	[0.019]	[0.025]	[0.019]	[0.025]	[0.018]	[0.024]
Stock returns [-61 to -2]	-0.007***	-0.005***	-0.007***	-0.004***	-0.007***	-0.004***	-0.007***	-0.004***	-0.007***	-0.004***
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
stdev(Stock returns [-61 to -2])	5.008***	-5.410***	4.064**	-10.656***	7.666***	-6.798***	7.034***	-12.038***	5.766***	-12.615***
	[1.393]	[1.848]	[1.588]	[2.217]	[1.426]	[1.999]	[1.597]	[2.333]	[1.557]	[2.248]
S&P500 returns [-61 to -2]	0.013***	0.017***	0.009***	0.003**	0.012***	0.016***	0.008***	0.003	0.009***	0.002
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
stdev(S&P500 returns [-61 to -2])	-0.018	-0.026	-0.025	0.173***	-0.044	-0.005	-0.069	0.192***	-0.059	0.194***
	[0.044]	[0.038]	[0.049]	[0.039]	[0.044]	[0.038]	[0.050]	[0.041]	[0.048]	[0.040]
Year dummies	No		Yes		No		Yes		Yes	
Industry dummies	No		No		Yes		Yes		Yes	
Observations	178898		178898		178898		178898		190807	
Pseudo R-squared	0.21		0.23		0.24		0.26		0.26	

Table 4: Firm and event characteristics by coverage probability decile

Panel A: Averages by absolute coverage (PMC) decile

	decile 1	*decile 2*	*decile 3*	*decile 4*	*decile 5*	*decile 6*	*decile 7*	*decile 8*	*decile 9*	*decile 10*
PMC	0.008413	0.017302	0.027334	0.039771	0.05629	0.079044	0.112622	0.16724	0.2703	0.527344
Earnings surprise	0.003717	-0.00032	-0.00212	-0.00237	-0.00293	-0.00388	-0.00405	-0.00469	-0.00501	-0.00338
Earnings (%)	-0.00524	-0.00251	0.001936	0.000777	0.000943	0.000186	0.003364	-0.00165	-0.00049	0.003925
Analyst attention	2.492233	3.027241	3.529878	4.125421	4.789786	5.632891	6.752067	8.322849	10.75336	14.55817
Market value	106886.6	206547.7	313184.5	454681.7	642810.9	901031.8	1379579	2234809	4301189	2.04E+07
$Trading volume	416132.9	1004774	1677842	2585093	3762556	5384633	8162267	1.35E+07	2.65E+07	9.54E+07
B/M (/1000)	0.003478	0.007449	0.005894	0.011443	0.028207	0.018266	0.021493	0.020908	0.022844	0.027833
Beta	0.515972	0.657666	0.729581	0.765989	0.802008	0.837004	0.86359	0.923508	0.95262	1.012684
Institutional ownership (%)	27.86756	35.44424	39.93365	43.33074	46.72175	49.53634	51.67275	54.58726	57.68943	58.00283
Recent returns	0.116278	0.097904	0.088688	0.079582	0.071903	0.0595	0.05621	0.04866	0.049333	0.050344
stdev(Recent returns)	0.034606	0.033447	0.032075	0.030665	0.029466	0.028437	0.027283	0.026306	0.024594	0.021796
CAR [-1 to 3]	0.676465	0.196768	-0.05085	-0.09165	-0.07879	0.024408	0.127856	-0.05123	0.063268	0.10237

Panel B: Averages by positive coverage (PMC^{+}) decile

	decile 1	*decile 2*	*decile 3*	*decile 4*	*decile 5*	*decile 6*	*decile 7*	*decile 8*	*decile 9*	*decile 10*
PMC^{+}	0.002361	0.006238	0.01131	0.018388	0.028197	0.042436	0.064677	0.102894	0.182529	0.410472
Earnings surprise	-0.01904	-0.00533	-0.0024	-0.00091	-9.3E-05	4.50E-06	0.000282	0.000421	0.000415	0.000278
Earnings (%)	-0.08633	-0.01245	0.004044	0.009397	0.012284	0.014246	0.015079	0.015418	0.012773	0.01064
Analyst attention	2.790695	3.271821	3.653985	4.230664	4.771599	5.487378	6.491396	8.021265	10.45043	14.44459
Market value	104382.9	197716	300878.6	436736.3	643220.1	919497.4	1386050	2199159	4059168	2.04E+07
$Trading volume	717795.8	1394021	2077001	3033687	4484215	6045202	8428405	1.32E+07	2.45E+07	9.28E+07
B/M (/1000)	0.007896	0.006333	0.007145	0.016943	0.024341	0.013944	0.018165	0.023894	0.011168	0.037184
Beta	0.679528	0.717723	0.737084	0.762249	0.780005	0.799122	0.815879	0.866857	0.911758	0.981299
Institutional ownership (%)	28.86881	34.01651	38.31315	42.45004	46.14008	49.60592	52.20011	54.83252	58.33189	58.2843
Recent returns	0.031128	0.060192	0.07302	0.078432	0.080355	0.0785	0.080836	0.078425	0.07621	0.077095
stdev(Recent returns)	0.04575	0.035586	0.032185	0.030013	0.027894	0.026498	0.024949	0.024113	0.022576	0.020481
CAR [-1 to 3]	-0.3868	-0.10396	-0.03111	0.032021	-0.01119	0.240443	0.221948	0.281683	0.341351	0.282786

Panel C: Averages by negative coverage (PMC^-) decile

	decile 1	*decile 2*	*decile 3*	*decile 4*	*decile 5*	*decile 6*	*decile 7*	*decile 8*	*decile 9*	*decile 10*
PMC^-	0.00289	0.006128	0.009698	0.014065	0.019719	0.027435	0.038257	0.054596	0.081806	0.178857
Earnings surprise	0.004752	0.001861	0.000766	-0.00016	-0.00144	-0.0022	-0.00317	-0.0041	-0.00615	-0.01584
Earnings (%)	0.01818	0.010028	0.002737	0.00098	0.003628	0.002024	-0.00062	-0.00348	-0.00674	-0.02656
Analyst attention	2.603615	3.23384	3.850433	4.462886	5.272751	6.247223	7.477658	8.733793	10.70223	11.82548
Market value	162139.9	308005.6	465044.8	671939.6	1009273	1710435	3234825	5160103	9273032	9484960
$Trading volume	644585.3	1465221	2390599	3585826	5261388	8320019	1.46E+07	2.32E+07	4.73E+07	5.44E+07
B/M (/1000)	0.00639	0.008788	0.010718	0.01534	0.015739	0.013679	0.029765	0.034414	0.016829	0.017235
Beta	0.513952	0.655139	0.708466	0.753413	0.789071	0.821657	0.875899	0.909814	0.971595	1.081843
Institutional ownership (%)	30.32216	38.06048	41.99662	45.00486	47.48159	49.20917	51.41433	53.14553	54.55903	54.90458
Recent returns	0.143468	0.121738	0.109492	0.090536	0.080553	0.065855	0.053646	0.047855	0.030247	-0.03035
stdev(Recent returns)	0.030146	0.030229	0.029934	0.029529	0.028772	0.028419	0.027824	0.027333	0.0268	0.029479
CAR [-1 to 3]	0.96151	0.581536	0.248212	0.125852	-0.0518	-0.02108	-0.02652	-0.05419	-0.18949	-0.69694

Table 5: Media coverage probability trading profits – Time-series tests

This table reports estimates of excess returns from OLS regressions of monthly PMC-portfolio returns (minus the risk-free rate) on a constant plus risk factors identified in the literature. Portfolios are formed based upon the most recent predicted probability of positive, negative, or no coverage in the preceding three months from the multinomial regression described in column 5 of Table 3 (omitting the year dummy variables). Portfolios are re-balanced monthly and returns are equal-weighted; * significant at 10%; ** significant at 5%; *** significant at 1%; Newey-West standard errors.

Panel A: Monthly portfolio alphas – CAPM model (MKT-Rf)

	Media coverage probability decile										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
Pr{Positive or Negative coverage | X}	0.00622**	0.00291	0.0013	0.00048	-0.0009	-0.0018	-0.00128	-0.00171	-0.00116	-0.00223***	0.00845***
	[0.00295]	[0.00240]	[0.00205]	[0.00196]	[0.00172]	[0.00171]	[0.00152]	[0.00141]	[0.00129]	[0.00083]	[0.00300]
Pr{Positive coverage | X}	-0.00392	-0.00023	0.00051	0.00019	0.00075	0.0014	0.00073	0.0004	0.00102	0.00076	-0.00468
	[0.00347]	[0.00244]	[0.00211]	[0.00192]	[0.00167]	[0.00168]	[0.00171]	[0.00150]	[0.00156]	[0.00076]	[0.00367]
Pr{Negative coverage | X}	0.00769***	0.00567**	0.00259	0.00124	0.00012	-0.00032	-0.00237	-0.00181	-0.00338**	-0.00770***	0.01539***
	[0.00287]	[0.00230]	[0.00194]	[0.00190]	[0.00183]	[0.00161]	[0.00151]	[0.00133]	[0.00135]	[0.00214]	[0.00378]

Panel B: Monthly portfolio alphas – Fama-French 3-factor model (MKT-Rf, SMB, HML)

	Media coverage probability decile										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
Pr{Positive or Negative coverage | X}	0.00443***	0.00193	0.0002	0.00001	-0.00148	-0.00265**	-0.00161	-0.00223*	-0.00183	-0.00254***	0.00697***
	[0.00165]	[0.00130]	[0.00105]	[0.00114]	[0.00103]	[0.00125]	[0.00131]	[0.00128]	[0.00126]	[0.00087]	[0.00164]
Pr{Positive coverage | X}	-0.00299	-0.00042	0.00047	-0.00076	-0.00025	0.00001	-0.00072	-0.00097	-0.00055	0.00024	-0.00323
	[0.00226]	[0.00182]	[0.00146]	[0.00126]	[0.00100]	[0.00100]	[0.00112]	[0.00099]	[0.00108]	[0.00070]	[0.00230]
Pr{Negative coverage | X}	0.00555***	0.00410***	0.00135	0	-0.00092	-0.001	-0.00277**	-0.00223**	-0.00320**	-0.00673***	0.01227***
	[0.00162]	[0.00121]	[0.00101]	[0.00105]	[0.00113]	[0.00091]	[0.00116]	[0.00113]	[0.00142]	[0.00228]	[0.00293]

Panel C: Monthly portfolio alphas – Carhart 4-factor model (MKT-Rf, SMB, HML, WML)

	Media coverage probability decile										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
Pr{Positive or Negative coverage | X}	0.00662***	0.00424***	0.00248**	0.00216*	0.00064	-0.00015	0.00081	0.0003	0.00023	-0.00035	0.00697***
	[0.00163]	[0.00115]	[0.00114]	[0.00113]	[0.00090]	[0.00090]	[0.00114]	[0.00123]	[0.00112]	[0.00082]	[0.00171]
Pr{Positive coverage | X}	0.00379	0.00399**	0.00319**	0.00151	0.00106	0.00117	0.0005	0.00016	0.00042	0.00113*	0.00266
	[0.00272]	[0.00197]	[0.00125]	[0.00095]	[0.00083]	[0.00087]	[0.00106]	[0.00097]	[0.00103]	[0.00067]	[0.00277]
Pr{Negative coverage | X}	0.00641***	0.00508***	0.00286***	0.00182	0.00095	0.00106	-0.00047	0.00033	-0.00002	-0.00111	0.00752***
	[0.00150]	[0.00115]	[0.00091]	[0.00111]	[0.00096]	[0.00105]	[0.00098]	[0.00093]	[0.00138]	[0.00213]	[0.00282]

Panel D: Monthly portfolio alphas – Pastor-Stambaugh 5-factor model (MKT-Rf, SMB, HML, WML, PS)

	Media coverage probability decile										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
Pr{Positive or Negative coverage | X}	0.00625***	0.00394***	0.00196	0.00174	0.00066	-0.00023	0.00007	-0.00027	-0.00064	-0.00036	0.00661***
	[0.00167]	[0.00126]	[0.00127]	[0.00121]	[0.00089]	[0.00101]	[0.00109]	[0.00112]	[0.00109]	[0.00081]	[0.00185]
Pr{Positive coverage | X}	0.00389	0.00369*	0.00280**	0.00147	0.00047	0.0011	-0.00029	-0.00078	-0.00035	0.00105*	0.00285
	[0.00281]	[0.00191]	[0.00124]	[0.00102]	[0.00091]	[0.00099]	[0.00120]	[0.00096]	[0.00103]	[0.00063]	[0.00291]
Pr{Negative coverage | X}	0.00585***	0.00486***	0.00237**	0.0017	0.00038	0.00069	-0.00117	0.00007	-0.00015	-0.00152	0.00738**
	[0.00152]	[0.00128]	[0.00106]	[0.00125]	[0.00106]	[0.00101]	[0.00106]	[0.00084]	[0.00143]	[0.00237]	[0.00323]

Table 6: Media coverage probability portfolios – Factor loadings (MKT-Rf, SMB, HML, WML)

This table reports estimates of excess returns from OLS regressions of monthly PMC-portfolio returns (minus the risk-free rate) on a constant plus MKT-Rf, SMB, HML, and WML. Portfolios are formed based upon the most recent predicted probability of positive, negative, or no coverage in the preceding three months from the multinomial regression described in column 5 of Table 3 (omitting the year dummy variables); Portfolios are re-balanced monthly and returns are equal-weighted; * significant at 10%; ** significant at 5%; *** significant at 1%; Newey-West standard errors.

Panel A: Portfolios formed on *Pr*{Positive or Negative coverage | X}$_{t-1}$

	Absolute (i.e. Positive or Negative) Coverage Probability Deciles										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
MKT-Rf	1.00392***	1.08844***	1.07511***	1.09574***	1.13996***	1.15116***	1.12191***	1.17927***	1.18616***	1.13669***	-0.13278***
	[0.04081]	[0.03274]	[0.02602]	[0.02285]	[0.02694]	[0.02706]	[0.02971]	[0.02450]	[0.03170]	[0.02054]	[0.04356]
SMB	0.89104***	0.90576***	0.89238***	0.80243***	0.71071***	0.62436***	0.56984***	0.42530***	0.31365***	0.05382**	0.83722***
	[0.07104]	[0.06555]	[0.06355]	[0.04530]	[0.05301]	[0.06345]	[0.04597]	[0.04790]	[0.04445]	[0.02729]	[0.06943]
HML	0.32118***	0.18504***	0.20392***	0.09423**	0.10812**	0.14231***	0.0516	0.07408*	0.09910**	0.02286	0.29833***
	[0.08418]	[0.05785]	[0.05721]	[0.04234]	[0.04448]	[0.04643]	[0.04404]	[0.03937]	[0.04105]	[0.02876]	[0.08594]
Momentum	-0.21324***	-0.22622***	-0.22272***	-0.21025***	-0.20681***	-0.24391***	-0.23633***	-0.24636***	-0.20027***	-0.21400***	0.00076
	[0.06179]	[0.03717]	[0.04544]	[0.03082]	[0.03027]	[0.02674]	[0.03809]	[0.04794]	[0.03645]	[0.02520]	[0.07111]
alpha	0.00662***	0.00424***	0.00248**	0.00216*	0.00064	-0.00015	0.00081	0.0003	0.00023	-0.00035	0.00697***
	[0.00163]	[0.00115]	[0.00114]	[0.00113]	[0.00090]	[0.00090]	[0.00114]	[0.00123]	[0.00112]	[0.00082]	[0.00171]
Observations	256	256	256	256	256	256	256	256	256	256	256
R-squared	0.84	0.91	0.93	0.94	0.95	0.94	0.94	0.94	0.94	0.96	0.53

Panel B: Portfolios formed on *Pr*{Positive coverage | X}$_{t-1}$

	Positive Coverage Probability Deciles										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
MKT-Rf	1.14264***	1.13650***	1.09520***	1.09535***	1.13327***	1.13535***	1.10124***	1.12067***	1.15004***	1.07008***	0.07255
	[0.05870]	[0.03744]	[0.04250]	[0.02636]	[0.02846]	[0.02696]	[0.02725]	[0.02122]	[0.02827]	[0.01850]	[0.06533]
SMB	1.29866***	1.00005***	0.91523***	0.79823***	0.67610***	0.55080***	0.45549***	0.34661***	0.20285***	-0.04702	1.34568***
	[0.07918]	[0.07033]	[0.05502]	[0.06183]	[0.05412]	[0.06470]	[0.06799]	[0.05289]	[0.05108]	[0.02936]	[0.08546]
HML	-0.18167*	0.02468	0.01973	0.17288***	0.18954***	0.24921***	0.25215***	0.23494***	0.26161***	0.07230**	-0.25396**
	[0.10872]	[0.07008]	[0.06129]	[0.04544]	[0.05750]	[0.06756]	[0.06279]	[0.04923]	[0.06196]	[0.03341]	[0.12669]
Momentum	-0.66268***	-0.43049***	-0.26536***	-0.22111***	-0.12803***	-0.11280***	-0.11937***	-0.11079***	-0.09455***	-0.08761***	-0.57507***
	[0.11655]	[0.09361]	[0.04832]	[0.03538]	[0.03180]	[0.03823]	[0.03688]	[0.03224]	[0.02708]	[0.02398]	[0.12477]
alpha	0.00379	0.00399**	0.00319**	0.00151	0.00106	0.00117	0.0005	0.00016	0.00042	0.00113*	0.00266
	[0.00272]	[0.00197]	[0.00125]	[0.00095]	[0.00083]	[0.00087]	[0.00106]	[0.00097]	[0.00103]	[0.00067]	[0.00277]
Observations	256	256	256	256	256	256	256	256	256	256	256
R-squared	0.84	0.89	0.93	0.94	0.95	0.93	0.92	0.93	0.92	0.96	0.68

Panel C: Portfolios formed on *Pr*{Negative coverage | X}$_{t-1}$

	Negative Coverage Probability Deciles										
	1 (low)	2	3	4	5	6	7	8	9	10 (high)	1 minus 10
MKT-Rf	0.99986***	1.06314***	1.09610***	1.10765***	1.08789***	1.10545***	1.11802***	1.12094***	1.16994***	1.31132***	-0.31146***
	[0.04295]	[0.02730]	[0.02535]	[0.02249]	[0.02836]	[0.02044]	[0.02979]	[0.02690]	[0.02905]	[0.05351]	[0.08613]
SMB	0.77203***	0.78715***	0.73100***	0.66560***	0.66586***	0.62559***	0.52462***	0.43760***	0.41571***	0.55980***	0.21223*
	[0.08121]	[0.06973]	[0.05154]	[0.07171]	[0.05220]	[0.04745]	[0.04880]	[0.04238]	[0.05278]	[0.06377]	[0.11111]
HML	0.39421***	0.29671***	0.22880***	0.22022***	0.18732***	0.12032***	0.06376	0.05711	-0.05322	-0.21614*	0.61035***
	[0.09377]	[0.08047]	[0.05910]	[0.05813]	[0.05311]	[0.03914]	[0.04687]	[0.03535]	[0.05498]	[0.11451]	[0.19054]
Momentum	-0.08458	-0.09590*	-0.14696***	-0.17769***	-0.18265***	-0.20137***	-0.22477***	-0.24974***	-0.31133***	-0.54903***	0.46445***
	[0.06440]	[0.04941]	[0.03767]	[0.03645]	[0.02677]	[0.03763]	[0.02843]	[0.03376]	[0.05907]	[0.10147]	[0.14250]
alpha	0.00641***	0.00508***	0.00286***	0.00182	0.00095	0.00106	-0.00047	0.00033	-0.00002	-0.00111	0.00752***
	[0.00150]	[0.00115]	[0.00091]	[0.00111]	[0.00096]	[0.00105]	[0.00098]	[0.00093]	[0.00138]	[0.00213]	[0.00282]
Observations	256	256	256	256	256	256	256	256	256	256	256
R-squared	0.84	0.9	0.93	0.94	0.94	0.94	0.94	0.95	0.93	0.89	0.43

Table 7: Media coverage probability portfolios – Double-sorts

This table reports estimates of excess returns from OLS regressions of monthly PMC-portfolio returns (minus the risk-free rate) on a constant plus MKT-Rf, SMB, HML, and WML. For each month, stocks are sorted first into quintiles based on firm or event characteristics, and then, within each of these quintiles, stocks are further sorted into quintiles based on lagged PMC. * significant at 10%; ** significant at 5%;*** significant at 1%; Newey-West standard errors.

Panel A: Portfolios formed within Market Value quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC-) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	0.00580***	0.00669***	0.00570***	0.00412	-0.00088	0.00668
		[0.00215]	[0.00213]	[0.00210]	[0.00329]	[0.00426]	[0.00406]
	2	0.00702***	0.00346**	0.00138	-0.00167	-0.00039	0.00741**
Market value		[0.00139]	[0.00154]	[0.00149]	[0.00137]	[0.00291]	[0.00343]
Quintiles	3	0.00593***	0.00286*	0.00085	0	-0.0029	0.00867***
		[0.00125]	[0.00146]	[0.00114]	[0.00115]	[0.00195]	[0.00250]
	4	0.00264*	0.00177	-0.00063	-0.00021	-0.00117	0.00381
		[0.00137]	[0.00129]	[0.00127]	[0.00122]	[0.00196]	[0.00249]
	(high)5	0.00149	0.00094	0.00057	0.00001	-0.00101	0.00233
		[0.00096]	[0.00123]	[0.00088]	[0.00083]	[0.00126]	[0.00190]

Panel B: Portfolios formed within Analyst attention quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC-) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	0.00710***	0.00599***	0.00444***	-0.00137	-0.00560**	0.01270***
		[0.00180]	[0.00117]	[0.00105]	[0.00176]	[0.00247]	[0.00305]
	2	0.00382**	0.00181	0.00287**	0.00048	-0.00268	0.00650**
Analyst attention		[0.00171]	[0.00125]	[0.00136]	[0.00178]	[0.00229]	[0.00289]
Quintiles	3	0.00508***	0.00354**	0.00182	0.00056	-0.00071	0.00579*
		[0.00161]	[0.00142]	[0.00121]	[0.00151]	[0.00248]	[0.00327]
	4	0.00343**	0.0002	0.00217*	-0.00055	-0.00039	0.00382
		[0.00154]	[0.00145]	[0.00117]	[0.00114]	[0.00237]	[0.00305]
	(high)5	0.00326*	0.00032	0.00244*	0.00226**	0.00105	0.00221
		[0.00176]	[0.00147]	[0.00125]	[0.00106]	[0.00218]	[0.00294]

Panel C: Portfolios formed within B/M quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC^-) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	0.00522***	0.00353	0.00015	-0.00002	0.00037	0.00485
		[0.00175]	[0.00252]	[0.00230]	[0.00246]	[0.00319]	[0.00352]
	2	0.00607***	0.00011	-0.00075	0.00114	-0.00044	0.00651***
B/M		[0.00191]	[0.00127]	[0.00130]	[0.00163]	[0.00145]	[0.00199]
Quintiles	3	0.00410**	0.00118	0.00101	-0.00190*	-0.00042	0.00527***
		[0.00170]	[0.00098]	[0.00119]	[0.00115]	[0.00119]	[0.00190]
	4	0.00674***	0.00229**	0.00059	0.00123	-0.00153	0.00827***
		[0.00166]	[0.00112]	[0.00124]	[0.00137]	[0.00149]	[0.00198]
	(high)5	0.00739***	0.00372**	0.00189	-0.00025	-0.00041	0.00806***
		[0.00149]	[0.00144]	[0.00162]	[0.00188]	[0.00235]	[0.00257]

Panel D: Portfolios formed within Beta (market) quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC^-) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	0.00604***	0.00243**	0.00052	-0.00384***	-0.00867***	0.01471***
		[0.00195]	[0.00102]	[0.00121]	[0.00124]	[0.00200]	[0.00225]
	2	0.00613***	0.00263**	0.00082	-0.00082	-0.00341***	0.00955***
Beta		[0.00162]	[0.00103]	[0.00132]	[0.00115]	[0.00122]	[0.00162]
Quintiles	3	0.00594***	0.00144	0.00118	-0.00059	-0.00149	0.00723***
		[0.00159]	[0.00144]	[0.00101]	[0.00118]	[0.00150]	[0.00171]
	4	0.00579***	0.00166	0.00299**	0.0016	-0.00033	0.00612**
		[0.00188]	[0.00156]	[0.00150]	[0.00158]	[0.00231]	[0.00236]
	(high)5	0.00756***	0.00404	0.00251	0.00458*	0.00305	0.00480*
		[0.00202]	[0.00247]	[0.00275]	[0.00271]	[0.00309]	[0.00274]

Panel E: Portfolios formed within Recent Returns quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC^-) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	0.00631***	0.00463*	0.00401*	0.00433	0.00184	0.00447
		[0.00186]	[0.00254]	[0.00240]	[0.00267]	[0.00335]	[0.00287]
	2	0.00473***	0.00513***	0.00205	0.00301*	0.00016	0.00456***
Recent Returns		[0.00155]	[0.00154]	[0.00143]	[0.00156]	[0.00110]	[0.00167]
Quintiles	3	0.00524***	0.00221	0.00143	-0.00008	-0.00131	0.00668***
		[0.00144]	[0.00141]	[0.00099]	[0.00140]	[0.00102]	[0.00184]
	4	0.00456***	0.00011	-0.00210*	-0.00298***	-0.00531***	0.00987***
		[0.00143]	[0.00110]	[0.00113]	[0.00083]	[0.00106]	[0.00201]
	(high)5	0.00804***	0.00424***	-0.00192	-0.00244	-0.00363*	0.01146***
		[0.00184]	[0.00138]	[0.00208]	[0.00197]	[0.00195]	[0.00285]

Panel F1: Portfolios formed within Earnings Surprise quintiles (Four-factor alphas)

		Probability of Negative Media Coverage (PMC⁻) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	-0.00428**	-0.00461***	-0.00208	-0.00143	-0.00355	-0.00074
		[0.00168]	[0.00169]	[0.00218]	[0.00202]	[0.00321]	[0.00344]
	2	-0.00097	-0.00190*	-0.00225**	-0.00232**	-0.00073	-0.00024
Earnings surprise		[0.00129]	[0.00100]	[0.00103]	[0.00113]	[0.00130]	[0.00199]
Quintiles	3	0.00874***	0.00509***	0.00368***	0.00171	0.00346**	0.00528*
		[0.00181]	[0.00133]	[0.00137]	[0.00122]	[0.00162]	[0.00275]
	4	0.00965***	0.00506***	0.00270*	0.0014	0.00096	0.00854***
		[0.00186]	[0.00126]	[0.00144]	[0.00157]	[0.00175]	[0.00236]
	(high)5	0.00987***	0.01116***	0.00767***	0.00377*	0.00186	0.00874***
		[0.00195]	[0.00175]	[0.00219]	[0.00213]	[0.00277]	[0.00287]

Panel F2: Portfolios formed within Earnings Surprise quintiles (Four-factor alphas)

		Probability of Media Coverage (PMC) Quintiles					
		1(low)	2	3	4	5(high)	1 minus 5
	(low)1	-0.00414**	-0.00365*	-0.00285	-0.00302	-0.00232	-0.00182
		[0.00188]	[0.00201]	[0.00180]	[0.00203]	[0.00286]	[0.00294]
	2	-0.00075	-0.00202**	-0.00225*	-0.00248**	-0.00067	-0.00008
Earnings Surprise		[0.00115]	[0.00095]	[0.00117]	[0.00108]	[0.00090]	[0.00139]
Quintiles	3	0.01002***	0.00413***	0.00279**	0.00205**	0.00395***	0.00606***
		[0.00173]	[0.00133]	[0.00136]	[0.00103]	[0.00101]	[0.00194]
	4	0.00832***	0.00642***	0.00255*	0.00143	0.00123	0.00716***
		[0.00165]	[0.00150]	[0.00147]	[0.00157]	[0.00121]	[0.00164]
	(high)5	0.01323***	0.00986***	0.00733***	0.00369**	0.00039	0.01264***
		[0.00210]	[0.00237]	[0.00214]	[0.00182]	[0.00158]	[0.00234]

Table 8: Cross-sectional regressions

Pooled OLS regressions of monthly stock returns minus the risk-free rate on the natural log of the previous month's value of PMC, lagged firm-event characteristics, and year and industry dummy variables; standard errors in brackets, robust to arbitrary heteroskedasticity and clustering by firm; * significant at 10%; ** significant at 5%;*** significant at 1%.

	$R_{i,t+1}$-r_f	$R_{i,t+1}$-r_f	$R_{i,t+1}$-r_f	$R_{i,t+1}$-r_f	$R_{i,t+1}$-r_f	$R_{i,t+1}$-r_f
log(PMC)	-0.00784***			-0.01534***		
	[0.00098]			[0.00136]		
log(PMC^+)		0.00306***			0.00156*	
		[0.00074]			[0.00091]	
log(PMC^-)			-0.00518***			-0.00690***
			[0.00056]			[0.00066]
log(Market value)	0.00417***	-0.00358***	0.00158***	0.00550***	-0.00178**	0.00071
	[0.00072]	[0.00060]	[0.00040]	[0.00084]	[0.00073]	[0.00064]
B/M (/1000)	0.00035	0.00034	0.00035	0.00038	0.00036	0.00036
	[0.00032]	[0.00031]	[0.00032]	[0.00033]	[0.00032]	[0.00033]
Beta [-61 to -2]	-0.00324***	-0.00378***	-0.00324***	-0.00374***	-0.00380***	-0.00360***
	[0.00074]	[0.00074]	[0.00074]	[0.00076]	[0.00076]	[0.00076]
Recent returns [-61 to -2]	-0.00019***	-0.00013***	-0.00018***	-0.00020***	-0.00013***	-0.00017***
	[0.00002]	[0.00002]	[0.00002]	[0.00002]	[0.00002]	[0.00002]
stdev(Recent returns [-61 to -2])	0.07166	0.0413	0.10907**	0.11255**	0.07215	0.14988***
	[0.04714]	[0.04788]	[0.04889]	[0.05114]	[0.05434]	[0.05285]
Institutional ownership (%)				0.00013***	0.00012***	0.00012***
				[0.00002]	[0.00002]	[0.00002]
log($Trading volume [-61 to -2])				-0.00002	-0.00196***	-0.00089
				[0.00057]	[0.00057]	[0.00055]
log(Analyst attention}				0.00966***	0.00265***	0.00654***
				[0.00091]	[0.00084]	[0.00081]
Earnings surprise				0.00028	0.04554**	0.01704
				[0.01936]	[0.01882]	[0.01908]
stdev(forecasts)				0.00082	-0.00063	0.00051
				[0.00055]	[0.00053]	[0.00054]
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180480	180480	180480	178144	178144	178144
R-squared	0.01	0.01	0.01	0.02	0.01	0.02

Table 9: Cross-sectional (Fama-MacBeth) regressions

Results are based on time series of estimated coefficients obtained from monthly regressions of returns on PMC values from the preceding month and lagged control variables; Fama-MacBeth standard errors in brackets; * significant at 10%; ** significant at 5%;*** significant at 1%.

Panel A

	Absolute Coverage Decile 1 dummy	*Absolute Coverage Decile 10 dummy*	*Positive Coverage Decile 1 dummy*	*Positive Coverage Decile 10 dummy*	*Negative Coverage Decile 1 dummy*	*Negative Coverage Decile 10 dummy*
Avg. coefficient	0.00638***	0.00002	-0.00031	0.00013	0.00632***	-0.00506**
F-M std. error	[0.00192]	[0.00199]	[0.00248]	[0.00194]	[0.00150]	[0.00203]
Number of monthly coefficients	257	257	257	257	257	257

Panel B

	log of Absolute Coverage Probability	*log of Positive Coverage Probability*	*log of Negative Coverage Probability*
Avg. coefficient	-0.01201*	-0.00091	-0.04254***
F-M std. error	[0.00673]	[0.00755]	[0.01242]
Number of monthly coefficients	257	257	257

Table 10: How does probability of media coverage predict stock returns?

Monthly four-factor alphas for decile portfolios formed on restricted versions of the PMC measures; each measure is orthogonalized with respect to cross-sectional estimates of size, B/M, beta, and momentum; Newey-West standard errors in brackets; * significant at 10%; ** significant at 5%;*** significant at 1%.

Panel A: PMC1 portfolio returns – expected media coverage conditional on the earnings surprise, the negative surprise dummy, the loss dummy, and the interaction between loss and negative surprise (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC1 (probability of any coverage)	0.00178	0.00800***	-0.00622***
	[0.00226]	[0.00209]	[0.00224]
$PMC1^+$ (probability of positive coverage)	-0.00217	0.00914***	-0.01131***
	[0.00182]	[0.00165]	[0.00222]
$PMC1^-$ (probability of negative coverage)	0.00870***	-0.00398*	0.01268***
	[0.00151]	[0.00215]	[0.00216]

Panel B: PMC2 portfolio returns – expected media coverage conditional on the seasonality, industry, and year dummies (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC2 (probability of any coverage)	0.00189	0.00042	0.00148
	[0.00133]	[0.00164]	[0.00179]
$PMC2^+$ (probability of positive coverage)	0.00105	0.00094	0.00011
	[0.00113]	[0.00169]	[0.00187]
$PMC2^-$ (probability of negative coverage)	0.00337**	-0.00019	0.00355
	[0.00166]	[0.00157]	[0.00223]

Panel C: PMC3 portfolio returns – expected media coverage conditional on the control dummies (PMC2) and the earnings information variables (PMC1) (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC3 (probability of any coverage)	0.00273*	0.00101	0.00172
	[0.00162]	[0.00139]	[0.00192]
$PMC3^+$ (probability of positive coverage)	-0.00177	0.00482***	-0.00659**
	[0.00214]	[0.00149]	[0.00266]
$PMC3^-$ (probability of negative coverage)	0.00620***	-0.0036	0.00980***
	[0.00162]	[0.00234]	[0.00282]

Panel D: PMC4 portfolio returns – expected media coverage conditional on: the earnings information variables (PMC1), plus log(analyst attention) and stdev(analyst forecasts) (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC4 (probability of any coverage)	0.00234**	0.00208	0.00026
	[0.00094]	[0.00233]	[0.00268]
$PMC4^{+}$ (probability of positive coverage)	-0.00134	0.00245	-0.00379
	[0.00181]	[0.00164]	[0.00251]
$PMC4^{-}$ (probability of negative coverage)	0.00559***	0.00055	0.00504*
	[0.00123]	[0.00246]	[0.00294]

Panel E: PMC5 portfolio returns – expected media coverage conditional on: the control dummies and earnings information variables (PMC3), plus log(analyst attention) and stdev(analyst forecasts) (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC5 (probability of any coverage)	0.00302***	0.00107	0.00195
	[0.00104]	[0.00199]	[0.00225]
$PMC5^{+}$ (probability of positive coverage)	-0.00085	0.00198	-0.00283
	[0.00157]	[0.00157]	[0.00227]
$PMC5^{-}$ (probability of negative coverage)	0.00513***	-0.00065	0.00578**
	[0.00128]	[0.00226]	[0.00274]

Panel F: PMC6 portfolio returns – expected media coverage conditional on: the control dummies, earnings information variables, and analyst variables (PMC5), plus value of trading volume, institutional ownership, and the U.S. firm dummy (orthogonalized with respect to size, B/M, beta, and momentum)

	four-factor alpha		
	PMC Decile 1 (low)	PMC Decile 10 (high)	PMC Decile 1 – Decile 10
PMC6 (probability of any coverage)	0.00515***	-0.00158	0.00673***
	[0.00131]	[0.00180]	[0.00195]
$PMC6^{+}$ (probability of positive coverage)	0.00185	0.00063	0.00121
	[0.00168]	[0.00151]	[0.00243]
$PMC6^{-}$ (probability of negative coverage)	0.00694***	-0.0033	0.01024***
	[0.00153]	[0.00223]	[0.00278]

Panel G: Long-range PMC5 portfolio returns
Month t+n four-factor alphas for long-short portfolios formed on the estimated media coverage probability observed in month t; expected media coverage conditional on: the control dummies and earnings information variables (PMC3), plus log(analyst attention) and stdev(analyst forecasts) (subsequently orthogonalized with respect to size, B/M, beta, and momentum); * significant at 10%; ** significant at 5%;*** significant at 1%.

	four-factor alpha		
Month	Any coverage (PMC5) portfolio Decile1-Decile10	Positive coverage ($PMC5^+$) portfolio Decile1-Decile10	Negative coverage ($PMC5^-$) portfolio Decile1-Decile10
t+1	0.002	-0.0028	0.00578**
	[0.00224]	[0.00226]	[0.00274]
t+2	0.00049	-0.00282	0.00364
	[0.00223]	[0.00243]	[0.00278]
t+3	0.00133	-0.00265	0.00359
	[0.00238]	[0.00243]	[0.00286]
t+4	0.00052	-0.00099	0.0009
	[0.00247]	[0.00242]	[0.00275]
t+5	0.00028	-0.00181	0.00159
	[0.00234]	[0.00241]	[0.00254]
t+6	0.00156	-0.00089	0.0004
	[0.00240]	[0.00237]	[0.00283]

Panel H: Long-range PMC6 portfolio returns
Month t+n four-factor alphas for long-short portfolios formed on the estimated media coverage probability observed in month t; expected media coverage conditional on: the control dummies, earnings information variables, and analyst variables (PMC5), plus value of trading volume, institutional ownership, and the U.S. firm dummy (subsequently orthogonalized with respect to size, B/M, beta, and momentum); * significant at 10%; ** significant at 5%;*** significant at 1%.

	four-factor alpha		
Month	Any coverage (PMC5) portfolio Decile1-Decile10	Positive coverage ($PMC5^+$) portfolio Decile1-Decile10	Negative coverage ($PMC5^-$) portfolio Decile1-Decile10
t+1	0.00676***	0.00124	0.01024***
	[0.00195]	[0.00243]	[0.00278]
t+2	0.00487**	0.00007	0.00624**
	[0.00200]	[0.00257]	[0.00294]
t+3	0.00722***	0.00288	0.00747**
	[0.00217]	[0.00259]	[0.00302]
t+4	0.00473**	0.00384	0.00582**
	[0.00188]	[0.00273]	[0.00281]
t+5	0.00429**	0.00392	0.00644**
	[0.00204]	[0.00278]	[0.00266]
t+6	0.00374*	0.00349	0.00442
	[0.00206]	[0.00253]	[0.00284]