News Consumption in the Wild*

J. Anthony Cookson† Colorado

Diego García ‡ Colorado

Elvis Jarnecic § University of Sydney

January 12, 2024

Abstract

We study how market returns shape news consumption, employing 700 million pageviews over 27 months from Australia's largest financial newspaper, the *Australian Financial Review*. Aggregate news consumption intensifies on days when the Australian market index decreases, led by a dramatic spike in consumption of markets news. By contrast, firm-specific news consumption declines when the aggregate market moves more (up or down). These findings imply aggregate and firm-specific news are substitutes for one another, consistent with theories of limited attention. These news consumption effects are strongest for fresh news, but they are also present for stale news articles on days when there are no articles about the firm.

Keywords: News; Information Consumption; Attention; Newspapers
JEL Codes: G12, G14, L82

*We thank George Wright, Lauren Vadnjal, and Agasthya Guruprasad at the Australian Financial Review for their partnership and for sharing the data. This paper has benefited from presentations at Baruch College, the Consumer Financial Decision Making lab group, London School of Economics, London Business School, CU Boulder Finance, and the Australian Financial Review. All remaining errors are our own. This work utilized the Alpine high performance computing resource at the University of Colorado Boulder. Alpine is jointly funded by the University of Colorado Boulder, the University of Colorado Anschutz, and Colorado State University. Data storage was supported by the University of Colorado Boulder 'PetaLibrary.'

†University of Colorado at Boulder. Email: tony.cookson@colorado.edu

‡University of Colorado at Boulder. Email: diego.garcia@colorado.edu

§University of Sydney. Email: elvis.jarnecic@sydney.edu.au

1 INTRODUCTION

Much of the finance literature is devoted to understanding how and why prices react to the arrival of news (Romer, 1992; Daniel, Hirshleifer, and Subrahmanyam, 1998; Cochrane, 2004). It is a standard view that news arrives at the moment it becomes available to the public (e.g., a news article posted online). However, for news to be impounded into market prices, it must first be consumed by investors. This perspective is well appreciated, though scholars typically infer how news is consumed indirectly, for example, by examining market reactions to the supply of news (Tetlock, 2007) or via events that highlight news arrival or distract from consuming it (Peress and Schmidt, 2020; Frydman and Wang, 2020). But little is known about what shapes news consumption because it is rare to directly observe *when* and *what* an investor reads, distinct from when the news article was supplied.[1] In this paper, we assemble a data set that provides precisely this information. Our empirical tests investigate how consumption of two types of news − aggregate market news and firm-specific news − relates to lagged stock market returns in the time series and in the cross-section.

Specifically, for a 27-month period from January 2021 to March 2023, our data contain all reader pageviews on articles posted to the online edition of the *Australian Financial Review* (AFR). AFR is the leading business newspaper in Australia and is similar to the *Wall Street Journal* in the United States. The data contain roughly 700 million pageviews on over half a million distinct pages, which is uniquely well suited to measure financial news consumption patterns. Each pageview encodes whether a user is browsing a specific article during a 10-second interval. Thus, the data give a complete picture of *when* news is consumed using timestamps, *what* news is consumed based on the characteristics of the articles themselves, and the *intensity* of news readership. Data on these aspects of news

[1]Similar to our data, Baba Yara, Davis, Grigoris, and Kantak (2023) use data on employee news consumption aggregated at the topic level to construct a novel measurement of uncertainty. Apart from addressing a different research question, our data environment is distinct in that we have fine-grained information on news supply and content.

consumption are unique to the literature, which typically observes when news is supplied and its content, but not the timing and intensity of news consumption.

Equipped with news consumption data, we start by examining how daily pageview counts relate to immediately preceding market returns. We conduct this analysis separately for pageviews before market opening, during market hours, and after market hours. It is important to highlight that this empirical design relies on three fairly independent experiments. Not only do the pageviews come from different non-overlapping parts of the day, but also, the lagged returns come from largely non-overlapping time windows (see Figure 1 for details). Across specifications, we find that news readership generally intensifies after large swings in the Australian market large cap index, the ASX 200. News consumption is especially sensitive to large market-wide losses, which is natural given the psychological bias to fixate upon negative outcomes (Baumeister, Bratslavsky, Finkenauer, and Vohs, 2001). The magnitudes we find are large and asymmetric: pageviews increase by 16% after the ASX 200 drops by one percentage point in the negative domain. In the positive domain, the increase is only 4%.

The overall pattern of news consumption following price movements is the average response, which potentially masks heterogeneity, as the full newspaper is a mixture of firm-specific, industry-specific, and market-wide commentary. Looking into heterogeneity, consumption of *general market* news increases sharply after the ASX index falls. By contrast, pageviews of *company-specific* news decrease after large swings in the ASX index. These findings suggest that jumps in the aggregate market draw news readers' attention away from firm-specific news and toward aggregate news.

Next, we analyze the cross-section of company news consumption, which complements the analysis of aggregate news consumption. In this analysis, firm-specific news consumption is responsive to recent returns, but in contrast to market news, the consumption of company news is more sensitive to positive returns than negative returns. We subject this cross-sectional result to a rich suite fixed effects that absorb unobservable, time invariant article and author characteristics that capture the nature of news supply (e.g., the *content* of the article). These tests exploit the fact that there are often multiple articles about the same firm that are *consumed* on the same day, which were sometimes authored on different days. In this article-firm-day panel, we find that the estimated magnitude is similar to our firm-day tests, which supports our interpretation that the responsiveness of news consumption to lagged returns reflects a change in demand for news that is unlikely due to a shift in supply.

In addition, we present two additional tests that distinguish news consumption from the supply of news. First, we restrict attention to firm-days with news consumption, but no news supply, by studying the subset of "stale" pageviews on articles that were published before day $t-1$. Even within this subset, firm stock price movements significantly predict news consumption. Second, we study news consumption of *pre-scheduled articles* for which the timing and content were decided on the prior day before the market signal is known. To do this, we identify articles that were published on *exactly* the top of the hour, which happens most commonly at 5 a.m (see Figure 4 for this notable spike in articles). These articles were typically written the day before, but they were embargoed from publication overnight with a scheduled release prior to the next day's news consumption. Consistent with market returns driving news demand, not supply, we see that these recent market returns predict news consumption of such embargoed articles.

Finally, in the cross-sectional specifications, we observe that firm-specific news consumption *goes down* when the aggregate market exhibits a larger magnitude change. This change $-$ which occurs in addition to the independent effect of firm-specific news $-$ suggests that aggregate news crowds out firm-specific news consumption. The findings suggest that not only are market news consumption and firm-specific news consumption distinct economically, but they also compete for readers' same attention resources.

Our main contribution is to the literature on the connection between market outcomes and news. This literature is core to asset pricing because it examines topics such as the impact of news (Busse and Green, 2002; Tetlock, 2007), the timing of market reactions and associated trading (Hirshleifer, Peng, and Wang, 2023), whether these impacts are permanent or reflect overreactions (De Bondt and Thaler, 1985; Jegadeesh and Titman, 1993; Daniel et al., 1998), and the extent of disagreement within and across news sources (Giannini, Irvine, and Shu, 2019; Cookson and Niessner, 2020; Fedyk, 2021). As the literature on media and markets has developed, scholars have provided deeper and more nuanced insight into the production of news (Gentzkow and Shapiro, 2010), particularly financial news (Dougal, Engelberg, García, and Parsons, 2012; Goldman, Gupta, and Israelsen, 2021), as well as when to expect news to have market impact (Martineau and Mondria, 2022). Our focus on news consumption is novel relative to this literature, which mostly learns from price reactions to announcements. For example, our evidence that firm-specific and aggregate news consumption are substitutes is novel and direct evidence in support of theories of limited attention and category learning (Peng and Xiong, 2006; Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016). This direct observation is useful as whether micro versus macro attention are complements or substitutes is not a settled question (Hirshleifer and Sheng, 2022).

Our research is also connected to the attention literature (Barber and Odean, 2008; Da, Engelberg, and Gao, 2011; Ben-Rephael, Da, and Israelsen, 2017; Ben-Rephael, Carlin, Da, and Israelsen, 2022), particularly the segment of the literature that examines attention to news media (Fedyk, 2022). Noting a paucity of data on information sources, recent research has turned to novel settings to identify information sources, such as account logins (Sicherman, Loewenstein, Seppi, and Utkus, 2016; Gargano and Rossi, 2018) and social media (Chen and Hwang, 2022; Cookson, Engelberg, and Mullins, 2023; Levy, Rossi, Shalev, and Zur, 2023; Cookson, Lu, Mullins, and Niessner, 2022). Our approach is to examine news consumption at its source, the newspaper, which draws a tighter connection to existing work on media. In this respect, our research is kindred to recent work by Kwan, Liu, and Matthies (2022) that links internet browsing data of firm-specific articles that originated from IPs of institutional traders, and Fedyk (2022) that studies how the positioning of Bloomberg articles affects their news readership. By contrast, our research studies the population of news readers of a major financial newspaper, which extends the analysis beyond professional and institutional investors. More instructively, we precisely delineate the timing and content of the news via our focus on a single newspaper.

Third, our research also connects to the literature on the consumption of stale news (Chang, Hartzmark, Solomon, and Soltes, 2017; Charles, 2022). Due to data limitations, this literature focuses on how *prices* of stocks respond to likely stale news consumption via news that is reprinted (Tetlock, 2011; Fedyk and Hodson, 2023). Although reposting the news is a natural shock to drive stale news consumption, it operates primarily through a *salience* channel (Frydman and Wang, 2020). The stale news effects we identify are conceptually distinct. Rather than reflecting salience of the individual article itself, we estimate that large swings in prices today correspond to when news readers revisit past articles about a firm. Such a mechanism is a natural part of the information gathering process, which is often slow to diffuse information into markets (Hong and Stein, 1999).

Finally, our research is relevant to the literature on the economics of media, particularly newspapers (Tetlock, 2007; García, 2013). Especially as newspapers face periods of prolonged decline (Gentzkow, Shapiro, and Sinkinson, 2014), leading to consolidation and technological changes (Ewens, Gupta, and Howell, 2022), it is instructive to understand the factors that intensify readership of newspapers. Our research highlights how firm-specific and market news respond to recent market conditions, showing that news readers' demand is driven by broad market patterns outside of editorial control. Such a finding suggests that newsrooms' effort to stay abreast of market conditions is an important aspect of catering to readers demand for news.

2 Data

In this section, we describe the underlying data in detail, and in so doing, provide motivation for the empirical tests.

2.1 Pageviews Data

We obtain our data from the Australian Financial Review (AFR), the leading business news daily in Australia, with a base of roughly 3.7 million average monthly readers (Financial Review, 2022). Our data consist of 700 million pageviews, spanning from January 1, 2021 through March 31, 2023. Each pageview encodes a 10-second interaction with AFR's online edition. These pageviews pertain to more than half a million distinct urls. Table 1 gives three examples of the pageviews in the data provided by AFR. For every pageview, we observe the precise timestamp, the title and url of the article, and importantly, the "referrer page," namely the url of the previous page *if* the reader came to the current page by clicking a hyperlink. In these examples provided the referrer is "internal," i.e. another AFR webpage, but we also see links from external sources (e.g., a Google or Facebook link). As we describe later, we use these referral patterns to reliably assign articles to firms without the need for a textual approach.

Though we also have some (masked) information on the identity of the user, we aggregate all the pageviews counts. We aggregate by url and date into pageviews "before market" (prior to 10am AEDT), "after market" (after 4 p.m. AEDT), and "during market hours" (between 10 a.m. and 4 p.m. AEDT). See Figure 1 for an illustration of our sample time frame. We note that, in contrast to previous work on newspapers, we have the precise timing of both news consumption and the news supply − i.e., the publication timestamp − for all articles available to the reader. We lever this data advantage by considering different within-day signals that could drive news consumption, i.e. stock returns the day before, returns the same day, etc.

Figure 2 plots the aggregate time series of all clicks in the AFR website during our sample period. The left panel refers to "before market" pageviews, the middle panel includes "during market" pageviews, and the last panel "after market" pageviews. The solid circles are pageviews on days where the Australian stock market is open. The blue triangles and red "x" refer to weekends and holidays, respectively. We observe approximately one million pageviews per trading day, with 40% of them occurring during market hours, the rest split evenly between before and after market hours. The volume of news consumption drops considerably on non-trading days, as well as around the turn of the year (late December and early January).

Figure 3 presents the daily total pageviews for six different articles, where the time starts on the day of publication of a story, running through its first 20 days of pageviews. Not surprisingly, we see a concentration of pageviews on the first day, when the article first makes it into the AFR website, with a fairly quick decay over the next few days. It is important to note that many articles do get "stale views," i.e. pageviews that occur days (even months) after publication. We use this variation in some of our tests, to rule out potential alternative interpretations.

2.2 Asset Data

The AFR website is structured as a standard online newspaper. It has a front page where the main news of the day are presented, with updates to the featured articles and sections throughout the day. The top of the front page, the most salient part of the AFR website, contains about ten articles, as well as links to the three most read columns of the newspaper: Street-Talk, which covers corporate topics (secondary offerings, M&A activity); Rear-Window, which discusses anything from politics to business, headed by Joe Aston since 2021; and Chanticleer, "a business column exploring the inner workings of individual companies, the executives who run them and their boards."

From this front page, readers can navigate to a typical news article, which we refer to as an "asset," or to a given section of the newspaper, where other news will be displayed. The main financial subsections of the AFR webpage include: *markets*, focused on general financial markets conditions; *companies*, where firm specific news are discussed; *property*, which covers real estate news. The AFR newspaper also has general interest sections outside business news, dealing with politics, world news, technology, etc.

Our analysis studies aggregate demand for news in a given section of the newspaper, i.e. pageviews in any url associated with the markets section of the newspaper, and also, firm- and article-specific demand, i.e. pageviews on a given piece of news on a given firm. The AFR primarily covers Australian firms, so our focus is on the largest publicly traded firms in Australia.

Table 2 gives aggregate statistics on the total number of pageviews per "section" of AFR, defined by the directory following the `www.afr.com` http address. The AFR front webpage, denoted as root in Table 2, received over 219 million pageviews during our sample. No article "resides" on this main page, but it is a substantial fraction of the pageviews in our sample. The root directory is not the only "non-asset" directory. Similar to the root directory pageviews, we see pageviews (but not asset pageviews) in the "topic," "company,", "search,", "by," and "markets-data" directories.

The most viewed section of AFR is `companies`, which has over 100 million pageviews. This section has 43,179 different assets, which received over 94 million total pageviews (6 million pageviews are to urls that do not correspond to an article, i.e. the root `companies` url and other subsections). Although more than 40,000 assets were viewed during our sample period, only 16,435 were originally published between 2021 and 2023. The final two columns indicate that the vast majority of the pageviews are on articles authored during our sample frame (84 million out of 94 million page views in the companies section). The next two most viewed business sections are `markets` and `property`. These three sections contain the bulk of the financial information produced by the AFR journalists, with over 65,000 assets across the three (more than 25,000 assets authored between 2021 and 2023), adding up to more than 150 million pageviews. The pageviews for Rear-Window and Chanticleer, the two columns highlighted above which are printed in the back page of the printed edition, have higher pageviews per asset, relative to other sections of the newspaper. We note that despite it being a business daily, there is significant consumption of general interest news, from political and policy issues to world news and personal finance stories.

In addition to the pageview data, AFR also provided metadata on the assets themselves, from the actual text of the article, to the author, as well as the date and time of publication, and the date the article was written. In Figure 4 we plot the histogram of the time of publication of the articles in the top panel. The publication times are uneven, reflecting the editorial practice to pre-schedule some articles for release at specific times. Notably, there is a large mass precisely at 5 a.m., which is the most common time for the editorial team to release an article that was written on a prior date. This uneven pattern of publication times (i.e., news supply) contrasts with the histogram of the timing of the pageviews (i.e., news demand), presented in the bottom panel of Figure 4. The highest demand for news is during the 9 a.m. hour, but news consumption is robust throughout the day from 8 a.m. through 5 p.m. There is non-trivial news demand both early in the day (6-8 a.m.) and in the late evening (5 p.m.-11 p.m.), but these periods have less than half of the news readership per hour compared to when the market is open.

2.3 Linking Assets to Firms

Each page url corresponds to an article or "asset." In addition to the main fields, we can see the referring article if a reader clicks on a link that brings them to such an article. A subset of these referring articles are static firm pages, which contain financial information, stock returns details, etc. These firm pages have urls of the form `https://www.afr.com/company/asx/jhx`, which identify the firm by its ticker (in this case, "jhx," see Figure A.1 for a screenshot). For the entire set of articles on AFR, we thus identify whether an article is about a particular firm if at least one user during our sample was referred to the firm's company page by a link from that article, or the reverse.

The two bottom examples in Table 1 illustrate this linking process. In the second entry, we have an article titled "AMPs stars fail to align, once again." This particular pageview occurred when a reader clicked on a embedded link on the company webpage for AMP (`/company/asx/amp`), which led the user to see the article in question (`/rear-window/amp-s-stars-fail-to-align-once-again-20220419-p5aecv`). The presence of this link and the reader's choice to click on it thus create direct linkage between the asset (p5aecv) and the firm ticker (amp). In the last entry of Table 1, we see the link in the other direction: the reader was browsing through an article (titled "James Hardie hikes prices twice in six months"), and then, clicked on a link that navigated to the company page for JHX immediately thereafter, creating a link between the asset (p5aluh) and the company (jhx).

This algorithm, at the asset-ticker level, only requires that *some* reader clicked on a referral link at *some* time during our data set. However, it allows us to reliably identify that an article is about a particular firm at the asset-ticker level for *all* pageviews, including pageviews from users who accessed the article via other navigation paths − e.g., from the front page, from a particular section, from a link on social media, or via direct navigation.

This algorithm gives us a list of 31,853 article-ticker pairs for 1,582 tickers. For our cross-sectional analysis, we restrict attention to stocks that are part of the ASX 200 index. Since some articles are associated with several stocks, we further restrict the article-ticker pairs to be sufficiently unique. In particular, out of all the ticker links associated to a given article, we only assign it to a given ticker if more than 90% of the links are associated with that one ticker. After these restrictions, we end up with a set of 14,342 AFR articles associated (uniquely) to 184 tickers.

2.4 Market Data and Timing

We collect daily price data, including the open, close and trading volume for all days and a broad sample of firms from the Securities Industry Research Centre of Asia Pacific, including the main stock index in Australia, the S&P/ASX 200, which we use in our aggregate time-series tests.

Our empirical specifications subdivide the day into "before market," "during market," and "after market" pageviews as separate dependent variables. To ensure a tight lead-lag relation between the returns variables and the pageview variables, we use close and open prices to construct within-day return variables for the return window that directly precedes each window in which we measure news consumption. Figure 1 presents the timeline of this measurement relative to the focal day t, and illustrates how we link each return variable to each news consumption variable in our subsequent empirical tests.

3 Results

In this section, we present several sets of results. We begin by examining how daily returns relate to the time series of daily aggregated news consumption. Next, we examine how returns predict news consumption in the cross-section, and we conclude the section with tests of mechanisms and the connection between aggregate market returns and firm-specific news consumption.

3.1 Market news

News consumption takes the form of count data on pageviews. Thus, we estimate a Poisson model, following Cohn, Liu, and Wardlaw (2022), using the following specification for the conditional expectation:[2]

$$\log(\mu(\text{News Consumption}_t)) = \gamma_t + \beta_1 \cdot \text{Returns}^+_{i,t-1} + \beta_2 \cdot \text{Returns}^-_{i,t-1} \tag{1}$$

where News Consumption_t is the number of pageviews of AFR readers on a particular section of the AFR website on date t, and γ_t denotes fixed effects (day-of-the-week and month).

For different specifications, we aggregate over a subset of the day's pageviews. In some instances, News Consumption_t counts pageviews separately based on the *time of day* – before 10am ("a.m. pageviews"), between 10am and 4pm ("mkt pageviews"), and after 4pm ("p.m. pageviews") – or the *newspaper section* – e.g., "markets," "company," "property." This specification links news consumption to *lagged* returns on the right hand side. To separately identify the impact of positive and negative returns, we employ two returns variables in each specification, Returns^+_{t-1} and Returns^+_{t-1}. The "+" version equals 0 for negative returns and equals the lagged return otherwise, whereas the "-" version equals 0 for positive returns but is equal to the lagged return otherwise.

When taking this specification to the data, the returns are computed from the ASX 200 Index's return based on the most recent within-day return window. For "before market" news consumption, Returns_{t-1} is the return from the close on date $t-2$ to the close on date $t-1$, the previous day index return. For "during market" news consumption, Returns_{t-1} is the return from the close on date $t-1$ and open on date t. For "after market" news consumption, Returns_{t-1} is the return from the close on date $t-1$ to the close on date t.

Table 3 presents the estimates from equation (1) for total news consumption across the full AFR webpage (any pageview), and pageviews of the front page (i.e., pageviews of "/") of the AFR website. From the first panel in Table 3, we estimate that before market pageviews are unrelated to lagged returns, on average. On the other hand, the

[2]The results in this section are similar if we use a log-linear specification and OLS, as the aggregate count data is fairly normal in log-space. For consistency with the rest of the analysis in the paper, where a Poisson model with fixed effects matters, we present this specification throughout.

close-to-open overnight return very strongly predicts news consumption during trading hours, particularly when the ASX index return is negative. The economic magnitude is large: a one percentage point drop in the ASX index in the negative domain is associated with 17.9% more pageviews of the AFR root directory. The corresponding estimate for returns in the positive domain is only 2.4%, not statistically different from zero. For after-market pageviews, we find a smaller effect in magnitude on the negative domain, but still statistically significant. We see a similar pattern for pageviews across all AFR articles, with similar economic magnitudes and statistical significance.

Table 4 presents a similar empirical design to that in Table 3, but instead, pageviews are aggregated by section. In these tests, the set of sections are the *markets* section, the *companies* section (see Section 2.2), and the *company* section (which we use to create asset-firm links in Section 2.3). The markets section is mostly about general market conditions, from stocks to currency movements and other macroeconomic events. The companies and company sections of AFR are firm specific news, i.e. firm events, CEO interviews, etc., and updates to financial information.

The first panel of Table 4 presents estimates based on counts of pageviews on the markets section. These estimates mirror those from the previous analysis from Table 3, but they exhibit a much stronger sensitivity of news consumption to lagged returns. We also observe a meaningful relation of news consumption to lagged returns in before market and after market periods. Specifically, a one percentage point increase in the positive domain in the ASX on day $t-1$ predicts a 4.4% increase in pageviews prior to the market open. By contrast, a one percentage point decrease in the ASX in the negative domain predicts 11.3% more before market news consumption. Both of these estimates are statistically different from zero and from each other. During trading hours, we see a much starker difference between the positive and negative domains. A percentage point drop in the ASX in the negative domain is associated with 56% more pageviews during trading hours, whereas the point estimate in the positive domain is only 1% and not significantly different from zero. When we examine news consumption after market hours (last two columns), we obtain similar estimates as when we examined before-market news consumption: a 19.8% change in the negative domain versus 8.4% in the positive domain.

One interpretation of these main findings is that the overnight returns in Australian markets indirectly proxy for return information from international markets, notably in the United States. As U.S. markets are open entirely during the Australian overnight period, the U.S. market signal is a potential alternative signal that could drive news consumption. To understand the importance of this alternative, in the Appendix, we enrich the specification with controls for the S&P500's daily returns. Appendix Table A.1 shows that our finding that the ASX index returns predict news consumption is not driven by U.S. market movements. That is, even holding constant the returns on the S&P500 index, movements in the ASX 200 index predict more news consumption, particularly in the negative domain and especially in the markets section.

Next, we examine the second and third panels of Table 4, which present evidence on the consumption of firm specific sections of AFR. The second panel presents estimates based on pageviews of the *company* section, which contains individual firms' company pages that contain financial information (see the discussion in Section 2.3). The third panel presents estimates based on pageviews of the *companies* section, which contains articles about individual firms. Surprisingly, in contrast to overall news consumption, news consumption on these firm-specific sections of AFR *goes down* when lagged returns swing by more. A one percentage point increase in the magnitude of negative returns predicts 7 percent less news consumption in the a.m. period, with a 9 percent reduction for a percentage point increase in returns in the positive domain. The numbers for during market hours mirror these numbers 17% (16%) decrease per percentage point change in ASX returns. And so do our estimates after market hours: 6.1% (6.5%) decrease per percentage point change in ASX returns. For pageviews in the companies section the results have the same signs, but the point estimates are smaller and more noisy (not statistically different from zero).

These results show that pageviews of company-specific information decrease as the market moves by more in magnitude, which is when AFR readers focus their attention on stories that are covered in the market section. Although these tests convey a clear message, the time series regressions are not causal evidence that returns drive demand for news. For example, it could be that AFR focuses on "general market" stories when the ASX has big swings, while releasing more firm-specific news when the ASX market is flat. For this reason, we turn to cross-sectional tests of firm-specific news consumption in the following section.

3.2 Firm-specific news

As with the aggregate news consumption, the firm-specific news consumption takes the form of count data on pageviews. Thus, we estimate a fixed effects Poisson model using the following specification for the conditional expectation:

$$\log(\mu(\text{News Consumption}_{i,t})) = \gamma_i + \gamma_t + \beta_1 \cdot \text{Returns}^+_{i,t-1} + \beta_2 \cdot \text{Returns}^-_{i,t-1} \tag{2}$$

where the dependent variable News Consumption$_t$ is the count of pageviews of AFR readers on articles linked to firm i on date t. In this firm-day panel setting, we follow closely the specification choices in the time series: we disaggregate daily counts of views by time of day $-$ before 10 a.m. ("before market hours"), between 10 a.m. and 4 p.m. ("during market hours"), and after 4 p.m. ("after market hours"). We restrict attention to articles linked firm i and we aggregate based on the timestamp of pageview t, not when the article was posted. As in the time series tests, the coefficients of interest are β_1 and β_2, which reflect how *lagged* returns predict news consumption, separately for when returns are positive (Returns$^+_{t-1}$) versus when returns are negative (Returns$^-_{t-1}$). In some specifications, we also include analogous returns variables for lagged returns on the ASX

200 index. The coefficients on these aggregate market terms capture how firm-specific news consumption is sensitive to movements in the aggregate market.

Following recent advances in efficiently estimating Poisson models (Correia, Guimarães, and Zylkin, 2020), we include firm and time fixed effects − either Month-year and Day-of-week fixed effects or Date fixed effects − across all specifications. For count data settings like ours, the recent applied econometrics literature has shown that Poisson fixed effects models are able to accommodate high dimensional fixed effects without any incidental parameters problem and can produce consistent estimates that are robust to misspecification (Lin and Wooldridge, 2019), in contrast to comparable OLS estimation of log transfomed counts (Cohn et al., 2022). We implement the fixed effects Poisson estimator using the `feglm()` function from the `fixest` library in R (Bergé, 2018), and we cluster standard errors by firm.

Table 5 presents the estimates from equation (2), with different choices of fixed effects. Across specifications, we find that firm-specific news consumption is significantly more responsive to lagged stock return movements in the positive domain than in the negative domain. Specifically, a one percentage point increase in returns in the positive domain is associated with 11 to 15 percent more news consumption. By contrast, we estimate that the same increase in firm-specific return magnitude in the negative domain (i.e., a one percentage point decrease in returns when returns are negative) predicts that news consumption only increases by between 3 and 6 percent.

Additionally, in the specifications that include the ASX index returns, we consistently estimate the *opposite* relation. That is, larger swings in the aggregate market − increases in the positive domain, decreases in the negative domain − are associated with less firm-specific news consumption. This finding reinforces our finding in the previous section that the "company" pages receive less visitation when the aggregate market swings by more. However, this test captures a broader class of firm-specific news consumption − articles that can be linked to specific firms, not just company pages − and the panel setting absorbs firm-specific unobservables and controls for monthly and day-of-week irregularities using fixed effects.

3.3 DISTINGUISHING NEWS CONSUMPTION FROM NEWS SUPPLY

In the previous section, we directly measure news consumption, not the supply of news. Although the measurement of news consumption is novel, an alternative interpretation is that readers' news consumption decisions are guided by the articles that were posted to AFR on that day. In this section, we develop two tests that disconnect the news consumption decision from the news supply decision.

In our first test, we explicitly focus on "stale" news consumption, or pageviews of articles that were not posted within at least a day of the article's publication. To do this, we restrict attention to pageviews of articles published before date $t-1$. In some specifications, we additionally drop firm-days in which no *other* article was published about firm i. We then analyze how counts of these stale pageviews depend on lagged returns in specifications perfectly analogous to equation (2), aggregated to the firm-day level for only stale pageviews. These estimates reflect how the consumption of news that was not authored on date t is consumed as a function of lagged returns. By disconnecting the timing of news supply from news consumption, the estimates arguably speak to how lagged returns drive news demand, absent differences in supply.

Table 6 presents the estimates using this measure of stale news consumption. Confirming the idea that news supply decisions matter for news consumption, the estimates are smaller for stale news consumption than overall. This decline in magnitudes is especially pronounced in the positive domain. Relative to the overall estimates of 11 to 15 percent, the coefficient estimates on positive returns are less than 6 percent for news consumption of stale articles. By contrast, the coefficient estimates on negative returns are more consistent in magnitude with the full sample: a one percentage point increase in the magnitude of negative returns predicts between 1.5 percent and 4 percent more news consumption of stale articles.

In our second test, we focus on news consumption of articles in which the supply decision was pre-scheduled, before the stock return signals are realized. To do this, we identify that were written distinctly before they were published but were embargoed for publication on a scheduled release. From discussions with AFR editorial staff, such embargoed articles are published exactly on the hour, most commonly at 5 a.m. AEDT. See the spike in publication times at 5 a.m. in Figure 4. The vast majority of these articles were written the prior day, but they were scheduled for delayed release. In the case of an article with a 5 a.m. release time, this implies that the article content and the decision to publish the article were decided upon *before* the overnight movement in returns is known. This information structure makes it unlikely that editorial decisions respond to market conditions, rather than news demand itself.

Table 7 presents the results from estimating (2), but only for pageviews on embargoed articles (i.e., those published exactly "on the hour"). Similar to our results on stale news consumption, we continue to estimate a strong asymmetry in how pre-market news consumption relates to lagged returns − a 7.8 to 9.2 percent increase for a one percentage point increase in returns in the positive domain (versus roughly 2 percent in the negative domain). We also see that that lagged returns predict during-market news consumption only in the negative domain, consistent with our stale news result. However, we do not see that evening pageviews are associated with lagged returns. Perhaps, this is due to most embargoed articles being published in the early morning hours.

Overall, the results in this section show that market returns predict news consumption, even for types of news consumption that are unlikely to be directly driven by news supply and editorial decisions.

3.4 ARTICLE-LEVEL NEWS CONSUMPTION

In this section, we consider an alternative data structure and specification that accounts flexibly for unobservable differences in the article content itself, as well as who authored it. To do this, we disaggregate counts of pageviews to the article-firm-date level. To focus on the days with the most meaningful news consumption, we restrict attention to the 10 days following the article's publication. This data structure allows us to include more granular fixed effects that can be linked to the article's characteristics (e.g., the author of the article or a fixed effect for the article itself).

To do this, we estimate a Poisson regression specification with fixed effects in the conditional mean, mirrioring the firm-date specification, as in:

$$\log(\mu(\text{News Consumption}_{a(i),t})) = FE(a,i,t) + \beta_1 \cdot \text{Returns}^{+}_{i,t-1} + \beta_2 \cdot \text{Returns}^{-}_{i,t-1} \quad (3)$$

where the dependent variable is News Consumption$_{a(i),t}$, which is the count of pageviews of AFR readers on date t for a specific article a that is linked to firm i. The variable definitions and implementation are perfectly analogous to the Poisson regression estimation the firm-day panel. We implement the fixed effects Poisson estimator using the `feglm()` function from the `fixest` library in R, but we do so on a more disaggregated data set with more refined fixed effects. Specifications include firm, article author, date, and in some cases article fixed effects. Consistent with the estimates in the firm-day panel, we cluster standard errors by firm.

Table 8 presents the results from estimating equation (3). Consistent with our main firm-day tests, we estimate that news consumption is typically more responsive to stock price movements in the positive domain (with exception of pm pageviews). We also estimate that both returns in the positive and negative domains explain significant variation in news consumption. Importantly, these specifications absorb time invariant differences in news consumption that are due to who authored the article (author fixed effects in the odd columns) or time invariant differences in news consumption by the article itself (article fixed effects in the even columns). Even within article, we find robust evidence that news consumption increases when a firm's return swings by more in the preceding period.

These specifications complement our analysis of stale news in that they offer an alternative way to hold constant the production of news. In so doing, our results speak to news consumption, distinct of the news supply decisions that have been widely studied in the literature.

4 Conclusion

Information consumption is central to the formation of asset prices (Ben-Rephael, Carlin, Da, and Israelsen, 2021). It is also critical to understanding the core business of news media. Despite its central importance, the literature has come to understand aspects of information consumption indirectly via market returns (Hirshleifer, Li, and Yu, 2015; Frydman and Wang, 2020) or via settings outside of traditional media (Cookson et al., 2023; Gargano and Rossi, 2018).

In this paper, we introduce novel data on information consumption from a major newspaper, the *Australian Financial Review*, spanning 27 months and consisting of 700 million pageviews. Our topline finding is that larger swings in market prices predict more consumption of financial news. This pattern is pervasive in that it is present in aggregate consumption of market news, but it is also evident in the consumption of firm-specific news in the cross-section. Interestingly, when we relate firm-specific news consumption to the aggregate market signal, we find that large swings in the aggregate market predict *reductions* in firm-specific news consumption. This finding suggests that aggregate and firm-specific news consumption are substitutes for one another, consistent with trading off information consumption under limited attention (Peng and Xiong, 2006; Kacperczyk et al., 2016).

More generally, our tests identify how news consumption responds to market conditions, distinct from the supply of news. Although we emphasize the role of recent market returns as a signal, we anticipate many fruitful paths to understand other drivers of news demand and their interaction with news supply by the media. We expect this agenda to become ever more important as the media landscape continues to change.

REFERENCES

Baba Yara, F., C. Davis, F. Grigoris, and P. Kantak (2023). Are uncertain firms riskier? *Carter and Grigoris, Fotis and Kantak, Preetesh, Are Uncertain Firms Riskier*.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies 21*(2), 785–818.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer, and K. D. Vohs (2001). Bad is stronger than good. *Review of General Psychology 5*(4), 323–370.

Ben-Rephael, A., B. Carlin, Z. Da, and R. D. Israelsen (2022). All in a day's work: What do we learn from analysts' bloomberg usage? *Available at SSRN 4056506*.

Ben-Rephael, A., B. I. Carlin, Z. Da, and R. D. Israelsen (2021). Information consumption and asset pricing. *The Journal of Finance 76*(1), 357–394.

Ben-Rephael, A., Z. Da, and R. D. Israelsen (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies 30*(9), 3009–3047.

Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers* (13).

Busse, J. A. and T. C. Green (2002). Market efficiency in real time. *Journal of Financial Economics 65*(3), 415–437.

Chang, T. Y., S. M. Hartzmark, D. H. Solomon, and E. F. Soltes (2017). Being surprised by the unsurprising: Earnings seasonality and stock returns. *The Review of Financial Studies 30*(1), 281–323.

Charles, C. (2022). Memory moves markets. *Available at SSRN 4019728*.

Chen, H. and B.-H. Hwang (2022). Listening in on investors' thoughts and conversations. *Journal of Financial Economics 145*(2), 426–444.

Cochrane, J. H. (2004). Asset pricing: Liquidity, trading, and asset prices. *NBER Reporter Online* (Winter 2004/05), 1–12.

Cohn, J. B., Z. Liu, and M. I. Wardlaw (2022). Count (and count-like) data in finance. *Journal of Financial Economics 146*(2), 529–551.

Cookson, J. A., J. E. Engelberg, and W. Mullins (2023). Echo chambers. *The Review of Financial Studies 36*(2), 450–500.

Cookson, J. A., R. Lu, W. Mullins, and M. Niessner (2022). The social signal. *Available at SSRN*.

Cookson, J. A. and M. Niessner (2020). Why don't we agree? evidence from a social network of investors. *The Journal of Finance 75*(1), 173–228.

Correia, S., P. Guimarães, and T. Zylkin (2020). Fast poisson estimation with high-dimensional fixed effects. *The Stata Journal 20*(1), 95–115.

Da, Z., J. Engelberg, and P. Gao (2011). In search of attention. *The Journal of Finance 66*(5), 1461–1499.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam (1998). Investor psychology and security market under-and overreactions. *the Journal of Finance 53*(6), 1839–1885.

De Bondt, W. F. and R. Thaler (1985). Does the stock market overreact? *The Journal of Finance 40*(3), 793–805.

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

Ewens, M., A. Gupta, and S. T. Howell (2022). Local journalism under private equity ownership. Technical report, National Bureau of Economic Research.

Fedyk, A. (2021). Disagreement after news: Gradual information diffusion or differences of opinion? *The Review of Asset Pricing Studies 11*(3), 465–501.

Fedyk, A. (2022). Front page news: The effect of news positioning on financial markets. *Journal of Finance, Forthcoming*.

Fedyk, A. and J. Hodson (2023). When can the market identify old news? *Journal of Financial Economics 149*(1), 92–113.

Financial Review (2022). Financial review increases readership despite softer post-covid market. Accessed on September 2, 2023.

Frydman, C. and B. Wang (2020). The impact of salience on investor behavior: Evidence from a natural experiment. *The Journal of Finance 75*(1), 229–276.

García, D. (2013). Sentiment during recessions. *The Journal of Finance 68*(3), 1267–1300.

Gargano, A. and A. G. Rossi (2018). Does it pay to pay attention? *The Review of Financial Studies 31*(12), 4595–4649.

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? evidence from us daily newspapers. *Econometrica 78*(1), 35–71.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2014). Competition and ideological diversity: Historical evidence from us newspapers. *American Economic Review 104*(10), 3073–3114.

Giannini, R., P. Irvine, and T. Shu (2019). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network. *Journal of Financial Markets 42*, 94–120.

Goldman, E., N. Gupta, and R. D. Israelsen (2021). Political polarization in financial news. *Available at SSRN 3537841*.

Hirshleifer, D., J. Li, and J. Yu (2015). Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics 76*, 87–106.

Hirshleifer, D., L. Peng, and Q. Wang (2023). News diffusion in social networks and stock market reactions. Technical report, National Bureau of Economic Research.

Hirshleifer, D. and J. Sheng (2022). Macro news and micro news: complements or substitutes? *Journal of Financial Economics 145*(3), 1006–1024.

Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance 54*(6), 2143–2184.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance 48*(1), 65–91.

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

Kwan, A., Y. Liu, and B. Matthies (2022). Institutional investor attention. *Available at SSRN 4073873*.

Levy, H., A. G. Rossi, R. Shalev, and E. Zur (2023). What's in investors' information set? *Available at SSRN 4068922*.

Lin, W. and J. M. Wooldridge (2019). Testing and correcting for endogeneity in nonlinear unobserved effects models. In *Panel data econometrics*, pp. 21–43. Elsevier.

Martineau, C. and J. Mondria (2022). News selection and asset pricing implications. *Available at SSRN 4194851*.

Peng, L. and W. Xiong (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics 80*(3), 563–602.

Peress, J. and D. Schmidt (2020). Glued to the tv: Distracted noise traders and stock market liquidity. *The Journal of Finance 75*(2), 1083–1133.

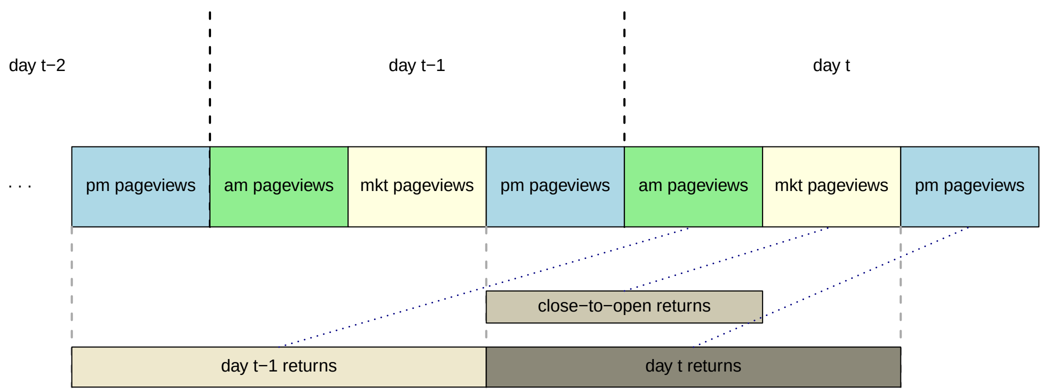
Romer, D. H. (1992). Rational asset price movements without news.

Sicherman, N., G. Loewenstein, D. J. Seppi, and S. P. Utkus (2016). Financial attention. *The Review of Financial Studies 29*(4), 863–897.

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

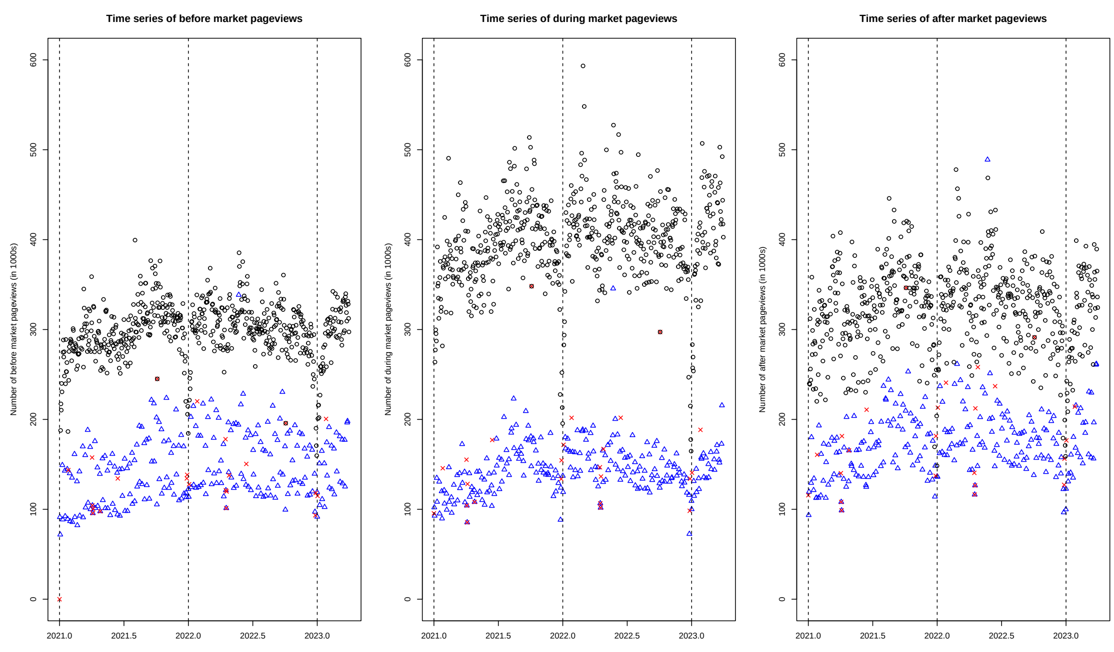
Tetlock, P. C. (2011). All the news that's fit to reprint: Do investors react to stale information? *The Review of Financial Studies 24*(5), 1481–1512.

Figure 1: Sample timeline of page views and returns measurement



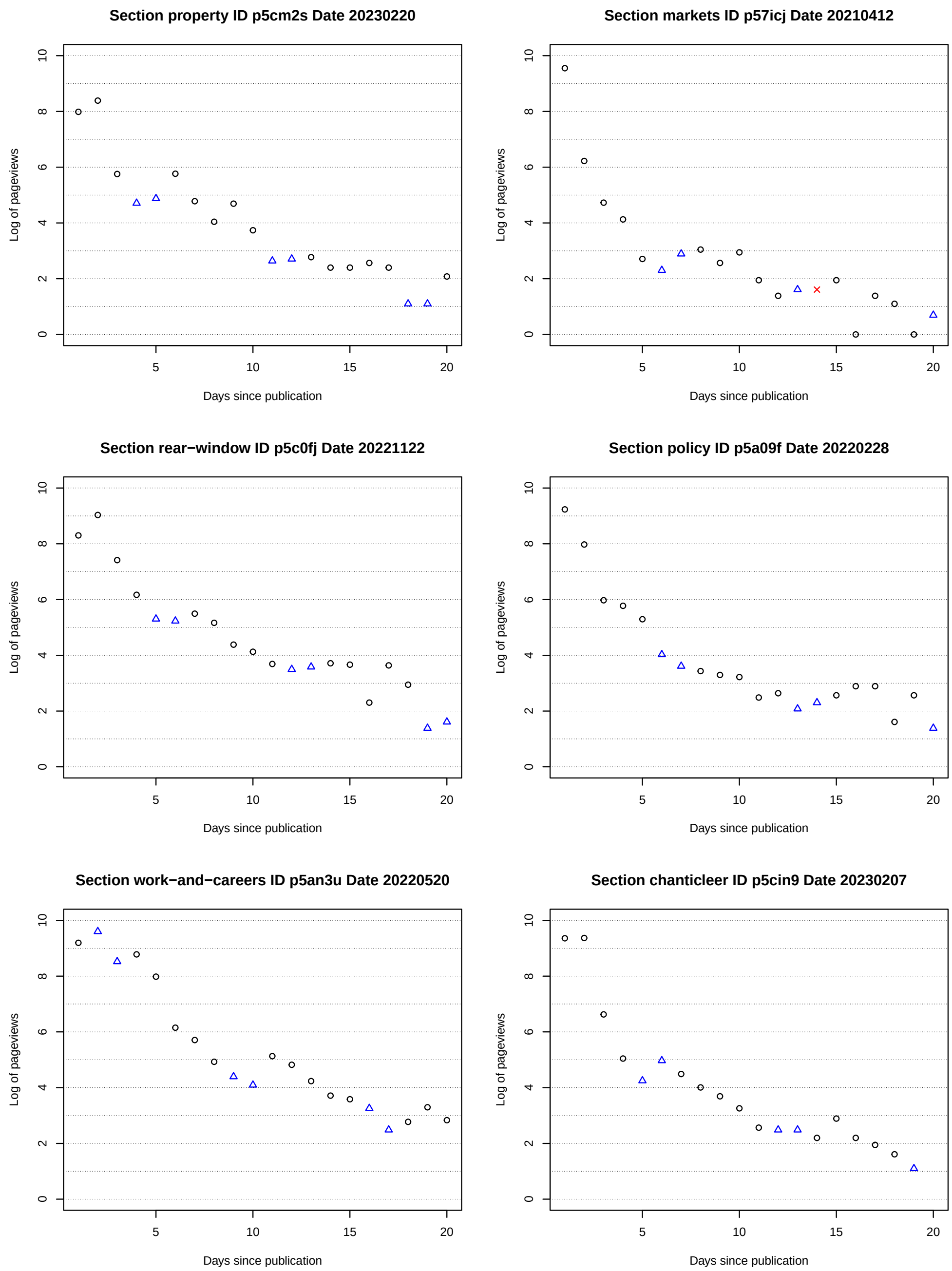
This figure illustrates measurement for our tests, which focus on page views within the focal day t. We divide each day into three based on Australian markets' opening time of 10 a.m. AEDT and closing time of 4 p.m. AEDT. Lagged return windows are illustrated in the bottom row, and the navy dotted lines show which returns variables we link to which news consumption variables.

Figure 2: Daily time series of aggregate page views on AFR



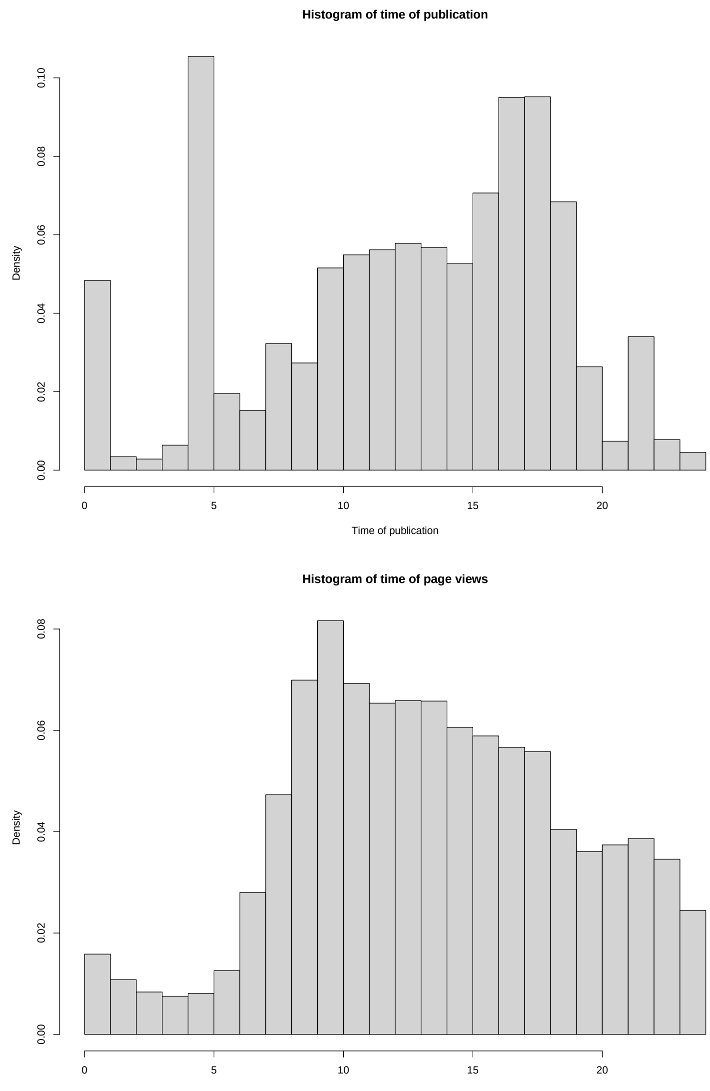
This figure reports the total number of pageviews on the AFR website during our sample period. Blue triangles refer to pageviews on weekends, red crosses pageviews on Australian holidays, and the solid circles are pageviews on trading days. The vertical dotted lines correspond to January 1 of each year in the data.

Figure 3: Examples of daily pageviews of AFR assets



This figure reports the total number of pageviews on six different stories in our dataset, during the first 20 days since publication. The y-axis uses a log scale. The title of each figure has the section of the article, its asset ID and date of publication. Blue triangles refer to pageviews on weekends, red crosses pageviews on Australian holidays, and the solid circles are pageviews on trading days.

Figure 4: Histogram of timestamps of publication and page views



The top panel plots a histogram of the time of publication of each article in our dataset. The bottom panel plots a histogram of the time of page views on trading days.

Table 1: Three sample pageviews

This table presents key information on three example pageviews that occurred on May 17th 2022. The timestamp is provided in UTC format.

Timestamp	2022-05-16 19:17:40.960
geo_timezone	Australia/Melbourne
page_title	ASX to rise, oil fuels Wall Street higher
page_urlpath	/markets/equity-markets/asx-to-rise-oil-fuels-wall-street-higher-20220517-p5alvu
page_referrer	/markets-data/world-indices
assetid	p5alvu
os_name	Android 5.x
Timestamp	2022-05-16 22:16:25.628
geo_timezone	Australia/Sydney
page_title	AMPs stars fail to align, once again
page_urlpath	/rear-window/amp-s-stars-fail-to-align-once-again-20220419-p5aecv
page_referrer	/company/asx/amp
assetid	p5aecv
os_name	Windows 8.1
Timestamp	2022-05-17 03:37:47.257
geo_timezone	Australia/Sydney
page_title	JHX News, Analysis, Announcements & Results ∣ James Hardie Industries Plc ∣ AFR
page_urlpath	/company/asx/jhx
page_referrer	/companies/infrastructure/james-hardie-hikes-prices-twice-in-six-months-20220516-p5aluh
assetid	NA
os_name	Ubuntu

Table 2: Descriptive Statistics

This table presents statistics on the number of pageviews (in millions) and the number of pages (i.e., unique urls) located in each of the main root directories of the online newspaper. The first two columns report pageviews across all urls on the website. The "All assets" columns count pages and pageviews for news articles with a publication date and contributed content (i.e., excluding "dynamic pages" like the front page or the root directory of the section). The last two columns ("Assets (2021−2023)") further restrict the counts of pages and pageviews to articles that are published during our observation period. The difference between the count of pageviews under "All assets" versus "Assets (2021−2023)" is the number of pageviews that occurred between 2021 and 2023 on articles that were published before 2021.

	urls		All assets		Assets (2021−2023)	
AFR root directory	Pageviews (in 1m)	# Pages	Pageviews (in 1m)	# Pages	Pageviews (in 1m)	# Pages
root	219.6	1	0.0	0	0.0	0
companies	100.3	43312	94.3	43179	84.5	16435
politics	55.0	19735	52.9	19703	49.9	7951
markets	41.3	9520	35.1	9454	34.1	5221
policy	37.2	14369	36.1	14338	33.9	6999
property	35.7	14951	32.0	14895	28.6	4382
street-talk	31.9	10261	21.4	10233	19.8	4714
world	30.3	11953	28.3	11937	27.2	7042
work	20.6	6182	19.4	6162	17.6	2615
technology	19.8	7756	18.8	7745	16.8	2936
wealth	19.6	4868	18.0	4846	16.6	2089
life-and-luxury	18.6	8880	17.4	8872	14.4	3180
rear-window	15.7	3513	15.2	3508	14.3	1285
chanticleer	15.0	3638	14.8	3632	14.2	1766
topic	10.2	1731	0.0	0	0.0	0
company	7.1	6166	0.0	0	0.0	0
search	5.8	1	0.0	0	0.0	0
by	2.3	983	0.0	0	0.0	0
rich-list	2.0	178	1.6	177	1.4	62
markets-data	1.7	38	0.0	0	0.0	0
opinion	1.3	1455	0.4	1453	0.0	33

Table 3: Aggregate News Consumption and Market Returns

This table presents Poisson regression estimates of the number of pageviews, predicted by preceding period stock returns. The returns variables are included: Returns^{-}_{t-1} for returns in the negative domain and Returns^{+}_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. The number of time-series observations in all regressions is $N = 567$ (all trading days in our sample).

Pageview window:	am		market		pm	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Root AFR						
Returns^{+}_{t-1}	−0.31	0.72	2.39	3.60	−0.66	0.72
Returns^{-}_{t-1}	−0.89	0.61	−17.89***	3.42	−1.98***	0.61
Pseudo R^2		0.347		0.493		0.753
All AFR						
Returns^{+}_{t-1}	−0.63	0.87	3.09	3.69	−0.73	0.98
Returns^{-}_{t-1}	−0.93	0.75	−14.92***	3.53	−2.24***	0.84
Pseudo R^2		0.313		0.411		0.557

Table 4: Aggregate Financial News Consumption and Market Returns

This tables presents Poisson regression estimates of the number of pageviews, predicted by preceding period stock returns. The returns variables are included: Returns^{-}_{t-1} for returns in the negative domain and Returns^{+}_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. The number of time-series observations in all regressions is $N = 567$ (all trading days in our sample).

Pageview window:	am		market		pm	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Markets section						
Returns^{+}_{t-1}	4.42^{**}	1.93	1.05	9.23	8.37^{***}	2.63
Returns^{-}_{t-1}	-11.34^{***}	1.56	-56.50^{***}	8.27	-19.82^{***}	2.08
Pseudo R^2		0.253		0.275		0.323
Company section						
Returns^{+}_{t-1}	-9.00^{***}	2.58	-17.49^{**}	6.83	-6.05^{***}	2.08
Returns^{-}_{t-1}	6.99^{***}	2.20	15.81^{**}	6.66	6.47^{***}	1.81
Pseudo R^2		0.097		0.160		0.142
Companies section						
Returns^{+}_{t-1}	-2.92	2.14	-0.32	8.60	-2.68	2.23
Returns^{-}_{t-1}	0.67	1.81	7.20	8.43	1.57	1.92
Pseudo R^2		0.162		0.206		0.272

Table 5: Firm-specific News Consumption

This tables presents fixed effect Poisson regression estimates of count of pageviews at the firm-day level, predicted by preceding firm-specific stock returns. The returns variables are included: Returns^-_{t-1} for returns in the negative domain and Returns^+_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. Standard errors are clustered by ticker and year-month. Firm, year-month, and day-of-week fixed effects are indicated at the bottom of the table. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Pageview window:	am		market		pm	
	(1)	(2)	(3)	(4)	(5)	(6)
Returns^+_{t-1}	11.2^{***}	11.7^{***}	13.6^{***}	15.2^{***}	11.3^{***}	13.0^{***}
	(0.91)	(0.98)	(1.07)	(1.19)	(0.95)	(1.12)
Returns^-_{t-1}	-3.01^{***}	-3.73^{***}	-3.23^{***}	-4.72^{*}	-3.47^{***}	-6.02^{**}
	(0.68)	(0.85)	(0.86)	(2.53)	(0.87)	(2.65)
ASX Returns^+_{t-1}	-9.26^{**}		-4.25		-8.63^{**}	
	(3.87)		(3.63)		(4.02)	
ASX Returns^-_{t-1}	3.44		6.96^{**}		8.60^{**}	
	(3.23)		(3.32)		(4.14)	
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Date		Yes		Yes		Yes
Year-month	Yes		Yes		Yes	
Day-of-week	Yes		Yes		Yes	
Fit statistics						
Observations	97,860	97,860	98,044	98,044	98,044	98,044
Squared Correlation	0.129	0.150	0.098	0.116	0.102	0.127
Pseudo R^2	0.391	0.411	0.384	0.405	0.408	0.431

Table 6: Firm-Specific News Consumption: Pageviews of Stale Articles Only

This table presents fixed effect Poisson regression estimates of count of pageviews at the firm-day level, predicted by lagged firm-specific stock returns. Only pageviews on articles published before date $t-1$ are included in these counts. In even columns, we additionally firm-days in which there is another AFR article posted about the firm. The returns variables are included: Returns^{-}_{t-1} for returns in the negative domain and Returns^{+}_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. Standard errors are clustered by ticker. Firm, year-month, and day-of-week fixed effects are indicated at the bottom of the table. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Pageview window:	am		market		pm	
	(1)	(2)	(3)	(4)	(5)	(6)
Returns^{+}_{t-1}	4.39***	2.80***	2.42**	0.067	7.02***	6.13***
	(0.92)	(0.99)	(0.99)	(1.10)	(1.21)	(1.29)
Returns^{-}_{t-1}	−1.99***	−1.55***	−2.21***	−1.98**	−5.14***	−4.01***
	(0.52)	(0.36)	(0.50)	(0.99)	(1.31)	(1.43)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Date	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	87,981	82,082	87,987	82,088	87,987	82,088
Squared Correlation	0.216	0.205	0.328	0.321	0.328	0.303
Pseudo R^2	0.526	0.505	0.596	0.577	0.615	0.594

Table 7: Firm-Specific News Consumption: Only Embargoed Articles

This table presents fixed effect Poisson regression estimates of count of pageviews at the firm-day level, predicted by lagged firm-specific stock returns. Only pageviews on articles published exactly "on the hour" are included in the analysis. These articles are typically embargoed − written at an earlier date but published at a planned future date and time. The returns variables are included: Returns^{-}_{t-1} for returns in the negative domain and Returns^{+}_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. Standard errors are clustered by ticker. Firm, year-month, and day-of-week fixed effects are indicated at the bottom of the table. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Pageview window:	am		market		pm	
	(1)	(2)	(3)	(4)	(5)	(6)
Returns^{+}_{t-1}	7.84***	9.24***	−0.69	−0.82	−5.04	−3.02
	(2.78)	(3.25)	(3.89)	(4.37)	(9.11)	(10.01)
Returns^{-}_{t-1}	−1.98***	−1.80***	−1.80**	−2.08***	−6.75	−6.11
	(0.61)	(0.65)	(0.72)	(0.74)	(4.38)	(6.10)
ASX Returns^{+}_{t-1}	−17.73**		29.61		−2.89	
	(8.96)		(26.8)		(12.51)	
ASX Returns^{-}_{t-1}	14.62**		57.11***		19.73	
	(7.26)		(21.01)		(14.80)	
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year-month	Yes		Yes		Yes	
Day-of-week	Yes		Yes		Yes	
Date		Yes		Yes		Yes
Fit statistics						
Observations	54,475	54,475	54,477	54,477	54,477	54,477
Squared Correlation	0.046	0.094	0.049	0.105	0.032	0.148
Pseudo R^2	0.271	0.343	0.288	0.360	0.319	0.411

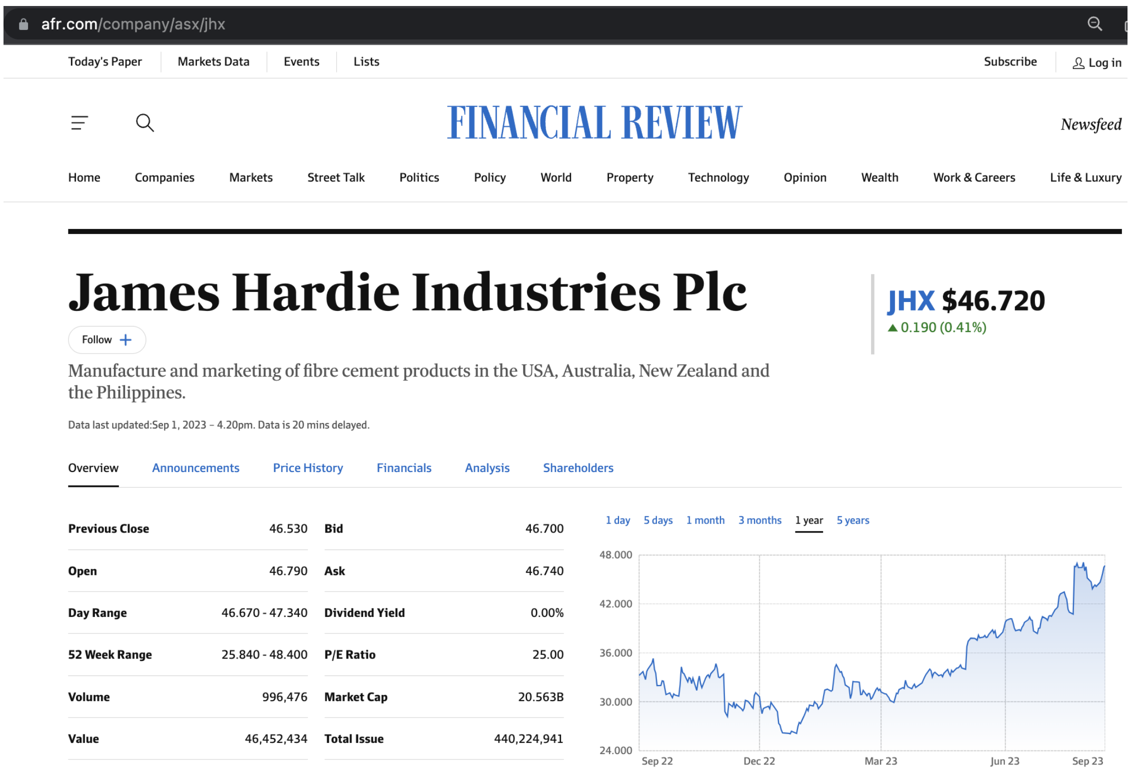
Table 8: Article-Level News Consumption and Firm-Specific Returns

This tables presents fixed effect Poisson regression estimates of count of pageviews at the asset-day level, predicted by preceding firm-specific stock returns. The sample is restricted to 10 days within the original article publication date. The returns variables are included: Returns^{-}_{t-1} for returns in the negative domain and Returns^{+}_{t-1} for returns in the positive domain. For "am" pageviews, the returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use return from close on date $t-1$ to open on date t. Standard errors are clustered by ticker and year-month. Firm, year-month, and day-of-week fixed effects are indicated at the bottom of the table. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

Pageview window:	am		market		pm	
	(1)	(2)	(3)	(4)	(5)	(6)
Returns^{+}_{t-1}	6.76***	7.75***	9.82***	16.30***	9.73***	13.90***
	(0.64)	(1.07)	(0.79)	(1.91)	(1.72)	(2.30)
Returns^{-}_{t-1}	−2.35***	−4.92***	−2.79***	−12.02***	−11.31***	−17.20***
	(0.52)	(1.69)	(0.80)	(2.66)	(1.70)	(1.97)
ASX Returns^{+}_{t-1}	−2.04	2.07	−27.51*	−44.50**	1.52	5.14
	(4.34)	(6.11)	(16.10)	(18.11)	(6.05)	(7.57)
ASX Returns^{-}_{t-1}	3.68	4.51	2.05	30.61	0.68	8.79
	(3.39)	(4.29)	(17.31)	(19.80)	(5.90)	(6.77)
Fixed-effects						
Firm	Yes	Yes	Yes	Yes	Yes	Yes
Year-month	Yes	Yes	Yes	Yes	Yes	Yes
Day-of-week	Yes	Yes	Yes	Yes	Yes	Yes
Author	Yes		Yes		Yes	
Article		Yes		Yes		Yes
Fit statistics						
Observations	63,507	63,504	63,536	63,534	63,536	63,536
Squared Correlation	0.010	0.197	0.009	0.231	0.010	0.244
Pseudo R^2	0.038	0.378	0.050	0.428	0.053	0.443

Internet Appendix

Figure A.1: An example of an AFR company page



This figure presents a screenshot of the AFR company page for James Hardie Industries Plc, available at https://www.afr.com/company/asx/jhx as an illustration of our company-asset linking process.

Table A.1: Aggregate Financial News Consumption and Market Returns

This table presents Poisson regression estimates of the number of pageviews, predicted by preceding period stock returns. The returns variables are included: ASX Returns$_{t-1}^{-}$ for ASX 200 index returns in the negative domain and ASX Returns$_{t-1}^{+}$ for ASX 200 index returns in the positive domain. We also include SPX Returns$_{t-1}^{+}$ and SPX Returns$_{t-1}^{-}$ for the date $t-1$ returns in the S&P500. For "am" pageviews, the ASX returns variables are close-to-close returns as of date $t-1$. For "pm" pageviews, the ASX returns variable is computed as the return from date t open to date t close. For clicks during the trading day, we use the ASX return from close on date $t-1$ to open on date t. The SPX returns are, for all windows, the returns on the SPX index during the trading day prior (the NYSE closes at 4pm ET, which corresponds to 4am-7am the following day in Sydney). *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively. The number of time-series observations in all regressions is $N = 567$ (all trading days in our sample).

Pageview window:	am		market		pm	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Root AFR						
ASX Returns$_{t-1}^{+}$	-0.52	0.75	0.27	4.11	-1.44^{*}	0.79
ASX Returns$_{t-1}^{-}$	-0.40	0.63	-16.85^{***}	3.99	-1.77^{**}	0.71
SPX Returns$_{t-1}^{+}$	0.93^{*}	0.51	0.62	0.71	1.28^{**}	0.55
SPX Returns$_{t-1}^{-}$	-1.21^{**}	0.48	-0.47	0.70	-0.56	0.55
Pseudo R^2		0.36		0.49		0.76
All AFR						
ASX Returns$_{t-1}^{+}$	-0.87	0.89	1.44	4.17	-1.93^{*}	1.08
ASX Returns$_{t-1}^{-}$	-0.62	0.75	-13.60^{***}	4.07	-2.35^{**}	0.97
SPX Returns$_{t-1}^{+}$	1.10^{*}	0.61	0.55	0.72	2.02^{***}	0.74
SPX Returns$_{t-1}^{-}$	-0.90	0.58	-0.56	0.71	-0.25	0.75
Pseudo R^2		0.32		0.41		0.56
Market						
ASX Returns$_{t-1}^{+}$	3.16^{*}	1.88	0.11	10.18	9.65^{***}	2.90
ASX Returns$_{t-1}^{-}$	-10.18^{***}	1.51	-42.71^{***}	9.42	-20.02^{***}	2.45
SPX Returns$_{t-1}^{+}$	4.81^{***}	1.29	1.92	1.75	-2.16	2.07
SPX Returns$_{t-1}^{-}$	-8.03^{***}	1.17	-5.16^{***}	1.64	0.66	1.98
Pseudo R^2		0.31		0.29		0.32