The Scandinavian
Journal of Economics

Scand. J. of Economics 124(3), 838–868, 2022
DOI: 10.1111/sjoe.12469

Asset returns, news topics, and media effects*

Vegard Høghaug Larsen†
Norges Bank, NO-0107 Oslo, Norway
vegard-hoghaug.larsen@norges-bank.no

Leif Anders Thorsrud
CAMP, BI Norwegian Business School, NO-0484 Oslo, Norway
leif.a.thorsrud@bi.no

Abstract

We decompose the textual data in a daily Norwegian business newspaper into news topics, and we investigate their predictive and causal role for asset prices. Our findings suggest that news published through the mass media has significant, persistent, and potentially economically profitable predictive power for returns. Moreover, during an exogenous media strike, returns for firms particularly exposed to our news measure experience a substantial fall relative to the control group. Together, these findings lend support for a view where the mass media acts as an "information intermediary" between agents and the state of the world, and disseminates fundamental information to investors.

Keywords: Latent Dirichlet allocation; machine learning; news; stock returns
JEL classification: *C*5; *C*8; *G*4; *G*12

1. Introduction

Can news in a business newspaper explain daily returns, and what is the effect of the media itself? As exemplified by Roll (1988), the economic literature has had a hard time finding a robust and positive answer to the first part of this question. One explanation often used for this rests on a view where arbitrage forces are unlimited and new information is incorporated into prices as soon as it is made public and before the (mass) media have time to report it. In contrast, using advances in natural language processing technologies, a small, but fast growing, body of

*This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. We thanktwo anonymous referees for comments that greatly improved this paper. We also thank Farooq Akram, Andre K. Anundsen, Drago Bergholt, Hilde C. Bjørnland, Jon Fiva, Hashem Pesaran, Johannes Skjeltorp, Rune Sorensen, Mike West, and colleagues at BI and Norges Bank for valuable comments. This work is part of the research activities at the Centre for Applied Macroeconomics and Commodity Prices (CAMP) at the BI Norwegian Business School.
†Also affiliated with CAMP at BI Norwegian Business School.

literature now tends to find a significant correlation between news and returns, suggesting that alternative theories about, for example, investor heterogeneity, sentiment, or rational attention are empirically relevant (Antweiler and Frank, 2006; Tetlock, 2007; Boudoukh et al., 2013; Peress, 2014; Calomiris and Mamaysky, 2017; Frank and Sanati, 2018). However, quantifying the potential causal link from the media to financial markets (i.e., answering the second part of the question above) is more difficult and much less explored. The difficulty arises because one has to separate the new information component from the effect of the news ether.

In this paper, we informally take the view that agents make endogenous information choices (Peng and Xiong, 2006; Kacperczyk et al., 2009; Schmidt, 2013) – but that no agent has the resources to monitor all the events potentially relevant for their decision – and thus delegate part of their information choice to specialized news providers, who report only a curated selection of events. As formalized in Nimark and Pitschner (2019), the media acts as an "information intermediary" between agents and the state of the world. This implies that editorial decisions can have independent effects on market developments.[1] We then operationalize this view by decomposing the textual information in a business newspaper into different types of news about economic developments, and we analyze market responses to these news items. Moreover, by utilizing an exogenous strike in the newspaper market, we argue that we are able to isolate the independent media component of the news signal from the new information component within this predictive relationship.

Our hypothesis is simple. To the extent that the newspaper provides a relevant description of the economy, the more intensively a given topic is represented in the newspaper at a given point in time, the more likely it is that this topic represents something of importance for the economy's current and future needs and developments. As such, it should also move stock prices. For example, we hypothesize that when the newspaper writes extensively about developments in, for example, the oil sector, and the tone is positive, it reflects that something is happening in this sector that potentially has positive economy-wide effects, and especially for firms related to the oil sector.

We apply our methodology to stocks listed on the Oslo Stock Exchange between 1996 and 2014, and use the entire newspaper corpus for Norway's major business newspaper, *Dagens Næringsliv* (DN). Although the Norwegian stock market is relatively small compared with international

[1]Rather than agents deciding *ex ante* on the expected usefulness of a particular signal, as in, for example, the costly information literature (Grossman and Stiglitz, 1980), knowledge of events is jointly determined *ex post* through a delegated information choice mechanism. See Larsen et al. (2021) for an application of this view in relation to inflation expectations.

markets, we focus on the case of Norway because it allows us to use a long history of the entire publications from the country's most important business newspaper, and we can use a well-defined exogenous strike in the Norwegian newspaper market, lasting for seven business days in 2002, to investigate the causal link from the media to financial markets.[2]

The newspaper content is available in the morning, at least two hours prior to when the market opens. Controlling for lagged returns, time- and firm-fixed effects, and other well-known predictors, numerous regressions show that a one unit positive innovation in the news predicts an increase of roughly 1 percentage point in close-to-open returns, and an increase of 1.5 percentage points in close-to-close returns. In the days following the initial news release, the effect accumulates further, suggesting a significant continuation pattern peaking at 4 percentage points after 15 business days, with little sign of reversal.[3] To gauge the robustness and economic significance of these pooled in-sample time series regressions, we implement simple out-of-sample zero-cost news-based investment strategies yielding significant annualized risk-adjusted returns ($Alpha$) of up to 20 percent.

During the strike period, the cross-sectional average return falls by roughly 60 basis points relative to the periods before and after the strike. However, this fall might be due to the strike effect itself, and not necessarily the media shortfall. Still, conditioning on how exposed the various firms are to our news measure during the year prior to the strike, we run a difference-in-difference identification strategy and find significant differences in mean returns of the same magnitude: returns for individual firms with a significant exposure to our news measures fall by 57 basis points during the strike period relative to firms with an insignificant news topic exposure. Because the average firm in the sample has a positive exposure to news, back-of-the-envelope calculations suggest that up to 20 percent of the predictive news–return effect can be attributed to the causal media effect.

These findings are important for two main reasons. First, they add further evidence to the empirical literature showing that news printed

[2] Up until 2002, DN was the leading (by far) business newspaper in Norway, making the Norwegian stock market an ideal candidate for our experiment. In other markets, where business news is distributed through a larger variety of sources, finding natural experiments, such as an exogenous strike that leads to a full media shortfall, would be much harder. We also note that DN is the fourth largest newspaper in Norway irrespective of subject matter.

[3] We also show that our results are not driven by well-known industry or day-of-the-week effects, and that they are not associated with firm characteristics such as book-to-market value, size, or liquidity. When analyzing the news–return relationship across three different subsamples, we do, however, find that the relationship becomes largely insignificant for the latter part of the sample (2008–2014). Interestingly, this loss of significance is alleviated when we expand the breadth of news sources utilized, suggesting that a broad-based news corpus needs to be applied to capture informative news signals in today's markets.

in mass media can explain daily returns. Although this finding is not new per se, we obtain it using a novel methodology not applied in the literature thus far, and test it on a new market. In particular, we use a latent Dirichlet allocation (LDA) model (Blei et al., 2003), proven to summarize textual data in much the same manner as humans would do (Chang et al., 2009), to decompose the textual information in a major business newspaper into news topics. Each topic is a distribution of words, and together the topics summarize the words and articles in the business newspaper into interpretable factors. We then construct series representing how much, and in which tone, each topic is written about in the newspaper across time, where news topics are tone adjusted (i.e., classified as either positive or negative news) using a dictionary-based approach commonly applied in the literature (see, e.g., Tetlock, 2007; Loughran and Mcdonald, 2011). Accordingly, the topic time series capture the continuously evolving narrative about economic conditions (Shiller, 2017).

Second, our study casts light on the various channels in which news (production) might affect returns. The fact that we find significant continuation patterns following news innovations, as opposed to reversal, suggests that our methodology parses out fundamental information, but also makes our results difficult to reconcile with a classical efficient market or investor sentiment view where prices should either respond immediately or overreact to news. Moreover, we find no significant differences in the return responses following either positive or negative news innovations. This makes it hard to explain our results using theories on investor heterogeneity, such as in Frank and Sanati (2018), where predictable patterns in returns are due to overreaction to good news and underreaction to bad news. Instead, we interpret our results more in line with the rational attention literature, and studies such as Peress (2014) and Li (2018), where the media helps to alleviate information frictions and to disseminate information to a large population of investors. We differentiate ourselves by potentially giving the media a larger independent role as an information intermediary in the news–return relationship. For example, during the strike, and given that our news items reflect positive information on average, a media information shortfall results in an adverse effect relative to the counterfactual.

The predictive part of our study is most closely related to studies by Antweiler and Frank (2006), Tetlock (2007), Boudoukh et al. (2013), and Calomiris and Mamaysky (2017). As in Tetlock (2007) and Boudoukh et al. (2013), we find an economically strong relationship between news and daily returns. Tetlock (2007) achieves this by constructing sentiment indicators from textual news using a dictionary-based method, and documents reversal patterns consistent with theoretical models of noise and liquidity traders. In contrast, we find significant continuation patterns following news innovations. Boudoukh et al. (2013) focus on news topics, and both their results and their approach are closer to ours. However, the methodology used in Boudoukh et al. (2013) relies on a substantial number of hard coded rules for classifying the news, while our approach utilizes a fully automated machine learning algorithm, which in principle is language agnostic. As such, our methodology is closer to those implemented in Antweiler and Frank (2006) and Calomiris and Mamaysky (2017), who use a naïve Bayes classifier and the Louvain method to derive news topics, respectively. We differ in that we do not limit ourselves to an event study approach, and we consider individual company returns on a daily frequency.[4]

In this respect, an additional novelty of our approach relates to how individual news topics are linked to companies using their textual description provided by Reuters. For example, a news topic that contains words mostly associated with the oil market will be linked to an oil company if the textual description of this company contains many of the same words. In contrast, typical textual approaches applied in the asset pricing literature link companies to (the derived) items in the news using explicit mentioning of their names, abbreviations, or other firm-specific characteristics. To us, this seems like an overly restrictive approach in as much as many news items might be relevant for stock prices without explicitly mentioning, for example, company names. Moreover, while such an identification scheme might work for larger economies and asset markets, it is less useful for smaller markets. The reason for this is that only a handful of companies are regularly explicitly mentioned in the mass media, and most companies would end up with having unrealistically few news days. In contrast, in our set-up, all days are news days, but to varying degrees. The validity of our approach for linking news to firms is tested when we randomly assign news topics to firms, and find that no significant predictive power between news and returns can be established in this case.

In terms of establishing a causal role of the media in financial markets, our study speaks directly to the studies by Engelberg and Parsons (2011), Dougal et al. (2012), Peress (2014), and Li (2018). Engelberg and Parsons (2011) document that news media coverage affects trading volumes. Dougal et al. (2012) appeal to a sentiment story whereby the bullish or bearish sentiment conveyed by newspaper columnists influence investors and returns. Peress (2014) and Li (2018) provide evidence more in line with ours, where the media alleviates information frictions and helps disseminate fundamental information to a large population of investors. In particular, whereas Li (2018) document that professional investors increase their information production and potential profits when media coverage

[4]Calomiris and Mamaysky (2017) use various decompositions of news articles to predict monthly and yearly risk and return developments in 51 aggregate stock markets. Antweiler and Frank (2006) run an event study covering US stocks and *Wall Street Journal* corporate news stories.

increases, we use the same exogenous strike as identified in Peress (2014) (for the Norwegian market) to disentangle the new information and media effect of the news signal. Novel to our study is that we apply a difference-in-difference identification scheme to control for the strike effect itself, and are able to provide a rough estimate of the relative importance of the media effect within a given predictive relationship.

The rest of this paper is organized as follows. In Section 2, we describe the data, the topic model, and how we link news to firms. In Section 3, we establish that news topics explain returns, while in Section 4 we investigate the causal impact of the media. We conclude in Section 5.

2. Data and news topics

The newspaper corpus used in this paper, the topic model specification, and the way in which news topics are transformed to time series follows Larsen and Thorsrud (2019) closely. We provide a summary of the computations below in Sections 2.1 and 2.2. In the interest of preserving space, technical details are delegated to Online Appendix B. New to this study is how we associate news topics to firms and returns. This is explained in Section 2.3.

2.1. The news corpus, the LDA, and topics

The DN news corpus is generously provided to us by the company Retriever through their "Atekst" database, and covers all articles published in DN from 2 May 1988 to 29 December 2014. In total, this amounts to 459,745 articles, well above one billion words, and more than a million unique tokens. This massive amount of data makes statistical computations challenging, but as is customary in the natural language processing literature, some steps are taken to clean and reduce the raw dataset before estimation. In particular, we remove stop-words, apply a stemming procedure, and reduce the number of unique words considered based on term frequency – inverse document frequency calculations. A description of how this is done is given in Online Appendix B.1. We note here that around 250,000 unique tokens are kept after the filtering procedure.

The "cleaned", but still unstructured, DN corpus is decomposed into news topics using an LDA model. The LDA model is an unsupervised topic model that clusters words into topics, which are distributions over words, while at the same time classifying articles as mixtures of topics. By unsupervised learning algorithm, we mean an algorithm that can learn/discover an underlying structure in the data without the algorithm being given any labeled samples to learn from. The term "latent" is used because the words, which are the observed data, are intended to communicate a latent structure, namely the meaning of the article. The term “Dirichlet” is used because the topic mixture is drawn from a conjugate Dirichlet prior. As such, the LDA shares many features with latent (Gaussian) factor models used in conventional econometrics, but with factors (representing topics) constrained to live in the simplex and fed through a multinomial likelihood at the observation equation. A richer description and more technical details of the LDA is provided in Online Appendix B.2. Here we note that we classify the DN corpus into $K = 80$ different topics using a Gibbs sampling algorithm. Although K might seem somewhat arbitrarily chosen, statistical tests conducted in Larsen and Thorsrud (2019) confirm that 80 topics give a good description of the corpus.

The LDA estimation procedure does not give the topics any name or label. To do so, labels are subjectively given to each topic based on the most important words associated with each topic. As shown in Table A1, in Online Appendix A, which lists all the estimated topics together with the most important words associated with each topic, it is, in most cases, conceptually simple to classify them. The labeling plays no material role in the experiment; it just serves as a convenient way of referring to the different topics instead of using, for example, topic numbers or long lists of words. What is more interesting, however, is whether the LDA decomposition gives a meaningful and easily interpretable topic classification of the DN newspaper. As illustrated in Figure A1, in Online Appendix A, it does. The topic decomposition reflects how DN structures its content, with distinct sections for particular themes, and that DN is a Norwegian newspaper writing about news of particular relevance for Norway. We observe, for example, separate topics for Norway’s immediate Nordic neighbors (“Nordic countries”); largest trading partners (“EU” and “Europe”); and biggest and second biggest exports (“Oil production” and “Fishing”). A richer discussion about this decomposition is provided in Larsen and Thorsrud (2019).

2.2. News topics as time series

Given knowledge of the topics (and their word distributions), the topic decompositions are translated into time series. This is done in two steps, which are described in greater detail in Online Appendices B.3 and B.4. In short, we first collapse all the articles in the newspaper for a particular day into one document, and compute, using the estimated word distribution for each topic, the topic frequencies for this newly formed document. This yields a set of K daily time series, where each represents how much (in percent) a given topic is written about for a given day. Then, for each observation in these time series, we identify their sign (i.e., whether or not the news is positive or negative). For each topic, this is done at the article level: for every daily observation, we find the article in the newspaper that is best explained by the topic. The tone of this article is identified using an external word list and simple word counts. The word list used here takes as a starting point the classification of positive/negative words defined by the *Harvard IV-4 Psychological Dictionary*, and then translates the words to Norwegian. The count procedure delivers two statistics, containing the number of positive and negative words. These statistics are then normalized such that each observation reflects the fraction of positive and negative words, and are subtracted from each other. If the difference is negative (positive), we set the sign equal to -1 (1), and adjust the topic frequencies accordingly.

We note that this procedure explicitly uses the output from the topic model also when defining the sign of the news, and that different topics might have their sign defined from the same article. Larsen (2021), Thorsrud (2020), and Larsen and Thorsrud (2019) have experimented with other ways of identifying the sign of the topic frequencies, finding that the method outlined above seems to work the best in a number of different applications.[5]

2.3. Financial data and linking news to firms

We obtain daily data for all firms listed on the Oslo Stock Exchange from Reuters Datastream. For each firm, we collect both open and close prices, and compute (log) close-to-open ($c2o$), open-to-close ($o2c$), and close-to-close ($c2c$) daily returns. We also collect the commonly used predictors (log) book-to-market (B/M), (log) market value (MV), and turnover ($Turn$), where the latter is computed by dividing the total number of shares traded by the number of shares outstanding. In addition, we use three measures of observed common time-fixed effects: (log) close-to-close returns on the Oslo Stock Exchange Benchmark index (R^{mh}), the close-to-close return on the S&P500 (R^{mi}), and the daily (log) change in the price of oil (R^{oil}).[6] Stocks listed for less than half a year are removed from the sample. To

[5]We have also used the word list suggested by Loughran and Mcdonald (2011) as a starting point for classifying positive/negative words, finding that this does not alter the end result by much. Still, there are undoubtedly more sophisticated methods that can be applied to identify the tone of the news (see, e.g., Pang et al., 2002).

[6]As roughly 50 percent of Norway's exports are linked to petroleum products and a large share of the companies traded on the Oslo Stock Exchange are directly exposed to the oil sector, controlling for the price of oil in asset pricing equations is often done when working with Norwegian data (see, e.g., Næs et al., 2009).

avoid including extreme price observations associated with listing and de-listing of firms, we exclude the first and last week of each firm's return observations. In total, we are left with 233 individual firms. The full sample stretches from 1996 to 2014, but only a few stocks are traded throughout the whole sample period.

To link companies to news, we use the word distributions estimated from the news corpus and each firm's textual description provided by Reuters. On average, across firms, the textual description is roughly a half-page description of what each company's primary business is. The firms' textual descriptions are then classified using a procedure for querying documents outside the set on which the LDA is estimated (Heinrich, 2009; Hansen et al., 2018). This corresponds to using the LDA model on the firm descriptions, but with the difference that the sampler is run with the estimated word distributions from the newspaper corpus held constant (see Online Appendix B.3). The end products of this procedure are vectors with topic probabilities for each firm description. From these vectors, we map firms with topics using the topic with the highest weight (probability) in describing the firm's core business.[7]

An example helps to illustrate our procedure. The first three, out of 10, sentences describing the firm *Frontline* are the following.

> Frontline Ltd. is a shipping company. The Company is engaged in the ownership and operation of oil tankers. The Company operates oil tankers of two sizes: very large crude carriers (VLCCs), which are between 200,000 and 320,000 deadweight tons, and Suezmax tankers, which are vessels between 120,000 and 170,000 deadweight tons ....

Following the steps described above, the elements with highest value in the vector with topic probabilities for this firm description are *Shipping*, *Airline industry*, and *Foreign*, with weights 0.31, 0.06, and 0.02, respectively. Thus, we associate *Frontline* with the *Shipping* topic. As seen from the word cloud for this topic (see Figure 1), the sentences and the word distribution for the topic share many important words. In our setting, the better the mapping is between the word distribution for a given topic and the words used in the description of the firm, the more likely it is that we match this particular firm with this topic.

[7] Mapping a firm to only one topic is simple and transparent, but still restrictive. However, when mapping firms to all topics using the topic weights, we observe that the predictive relationship between news and returns described in later sections tends to become insignificant. One likely reason for this is a strong deterioration of the signal-to-noise ratio. We leave it for future research to explore more sophisticated mapping alternatives.

Figure 1. The *Shipping* topic represented as a word cloud

Notes: The size of a word reflects the probability of this word occurring in the topic. The word cloud is created based on the 40 most important words in the topic.

Summary statistics for our firm and topic mapping are provided in Table A2 and Figure A2 in Online Appendix A. We highlight four overall impressions. First, of the 80 individual news topics, 33 are successfully mapped to one or more firms. We have informally looked at all estimated firm and topic mappings. It is our impression that the procedure produces intuitive mappings in well over 70 percent of the cases. However, for some firms, the topic mapping seems weird. While we could have excluded companies from the sample in such cases, or manually changed the mapping, we have refrained from doing so to keep the analysis as transparent as possible. Second, a large share of topics are mapped to relatively few companies. However, a large share of the firms are linked to the topics *Oil service*, *Shipping*, and *IT systems*, in particular. Third, there are large variations in average returns, standard deviations, average size and book-to-market values across firms with different topic mappings, but no clear patterns indicating a systematic relationship. Finally, we see from Figure A2 that the numbers of firms included in the panel dataset, and which topic assignments that dominate, change considerably across time. For example, during the 1990s, the breadth of the market was much thinner than during the 2000s, and topics associated with oil, health, and fishery have become more important over the years.

3. News and returns

Let $y_{i,t}$ denote the return series for company i at time t, where we consider both the (log) change in the price from market closing on day $t-1$ to either the opening ($c2o$) or closing price ($c2c$) on day t as return measures. Then, our main statistical tool for analyzing whether or not news explains returns is a general panel data regression of the form,

Table 1. Firm-specific news topics and day t close-to-open and close-to-close returns

	Close-to-open returns			Close-to-close returns			Std(X)
	(I)	(II)	(III)	(I)	(II)	(III)	
$Topic_t$	0.0107***	0.0135***	0.0090**	0.0153***	0.0272***	0.0208***	0.017
	(0.0022)	(0.0031)	(0.0037)	(0.0031)	(0.0051)	(0.0062)	
R^{mi}_{t-1}		0.3453***	0.3410***		0.2816***	0.2774***	0.013
		(0.0177)	(0.0217)		(0.0285)	(0.0371)	
R^{mh}_{t-1}		−0.0120	−0.0180		−0.0012	−0.0023	0.016
		(0.0133)	(0.0158)		(0.0222)	(0.0275)	
R^{oil}_{t-1}		0.0139**	0.0067		0.0231*	0.0132	0.020
		(0.0067)	(0.0091)		(0.0120)	(0.0169)	
B/M_{t-1}		−0.0007***	−0.0007***		−0.0008***	−0.0007***	0.889
		(0.0001)	(0.0001)		(0.0002)	(0.0002)	
MV_{t-1}		0.0002*	0.0002*		−0.0003*	−0.0003*	1.802
		(0.0001)	(0.0001)		(0.0001)	(0.0002)	
$Turn_{t-1}$			0.0115***			0.0076***	0.047
			(0.0021)			(0.0022)	
R^2	0.0088	0.0298	0.0301	0.0054	0.0128	0.0128	
Obs.	540,708	540,708	391,533	540,708	540,708	391,533	
α_i	Yes	Yes	Yes	Yes	Yes	Yes	
δ_t	Yes	No	No	Yes	No	No	

Notes: In each regression, the key independent variable is $Topic_t$. All regressions control for the firm's lagged close-to-close return (R_{t-p}), for $p = 1, \ldots, 14$. Control variables listed in the table include: close-to-close return on the S&P500 (R^{mi}_{t-1}), close-to-close returns on the OSEBX (R^{mh}_{t-1}), the daily change in the oil price (R^{oil}_{t-1}), the book-to-market value (B/M_{t-1}) the market value (MV_{t-1}), and finally the turnover ($Turn_{t-1}$). All regressions are estimated by OLS. Fixed effects are included as specified in the table. Following Tetlock et al. (2008), we compute clustered standard errors by trading day. Robust t-statistics are given in parentheses. The last column reports the unconditional standard deviation of the individual predictors. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

$$y_{i,t} = \gamma T_{k,t} + z'_{i,t}\boldsymbol{\delta} + \alpha_i + \delta_t + u_{i,t}, \tag{1}$$

where $T_{k,t}$ is the news topic, associated with firm i, and γ is the parameter of interest. The newspaper, and thus $T_{k,t}$, becomes available after the market closes on day $t-1$, and usually in the morning on day t, at least two hours prior to when the market opens. Accordingly, the timing used in equation (1) ensures that we do not use news that was generated by market movements on day t itself. Additional commonly used predictors are included in the vector $z_{i,t}$, and α_i and δ_t are firm- and time-fixed effects, respectively.

Table 1 highlights our first main result. Regressing $c2o$ returns on the news measure produces positive and highly significant news coefficients. Controlling only for lagged returns ($c2c$) and firm- and time-fixed effects, we see from Column I that a one unit (standard deviation) innovation in the news corresponds to roughly a 1 (0.02) percent increase in returns. Augmenting the regressions with various other control variables does not alter this finding by much. At most, we obtain an effect size of 1.35 percent, and, when controlling for turnover, in Column III of the table, the effect size is 0.9 percent. Note here, however, that the number of observations is somewhat reduced as not all firms have recorded turnover for the whole sample period. As is common in this type of regressions, the R^2 is low, indicating that most of the day-to-day variation in individual firm valuations is idiosyncratic. We observe from Columns II and III, however, that the returns on the previous day's S&P500 (R^{mi}_{t-1}) have a particularly positive and strong predictive power for the subsequent returns. In unreported results, we also confirm that it is this variable that attributes the most to the increase in R^2 across Columns I and II. The US market closes over six hours after the Norwegian market, so the R^{mi}_{t-1} variable also contains more timely information than any of the other variables used in the regressions. Still, including the S&P500 hardly changes the size and significance of the news coefficient. Among the other potential determinants of $c2o$ returns, the (log) change in oil prices (R^{oil}_{t-1}), book-to-market (B/M_{t-1}), market-value (MV_{t-1}), and turnover ($Turn_{t-1}$) all show signs of being significant, confirming well-known asset pricing results.[8]

Because of the short window between when the newspaper is released in the morning and when the market opens, it is natural to interpret the findings thus far as saying that the news topic variables capture new information that the market responds to. This is not to say that it is the newspaper that generates this news. For example, firm-specific news might be released after the market closes on day $t-1$, and then written about in the newspaper that is published in the morning on day t. Still, although the media might report on already known information, the fact that they actually report on it, and the intensity and manner in which they do so, might have a separate effect on asset pricing valuations. We investigate this further in Section 4. Below, however, we first investigate whether the news topics carries fundamental information or noise, and we report on various robustness checks, including assessments of particular time periods, and implement a simple trading strategy.

3.1. Continuation or reversal?

A classical finding in finance is that investors overreact to noisy information, and underreact to new fundamental information (see, e.g., French and Roll,

[8] Petersen (2009) documents how previous results in the asset pricing literature are highly sensitive to how the standard errors in panel data regressions are computed. In unreported results, we show that all of our significant tests are robust to clustering the standard errors on either time, firms, and groups (topics), i.e., t, i, and k using the notation from equation (1). Irrespective of clustering level, the news coefficients are always significant at either the 1 or 5 percent level.

1986; Campbell et al., 1993). This results in significant continuation patters in returns following new information about fundamentals, but a reversal following information that turned out to be noise.

The columns for close-to-close returns in Table 1 reproduce the regressions discussed above, but now using close-to-close returns as the dependent variable. Accordingly, compared with using $c2o$ returns, prices have a longer time to respond to the news signal (which is released early in the morning on day t). As seen from the table, the news variable remains highly significant, but the magnitude of the effect is somewhat larger than previously found. Now a one unit positive news innovation translates into a 1.53 percent increase in returns for the specification reported in Column I, and up to 2.72 percent for the specifications reported in Column II. These numbers are approximately 50 and 130 basis points larger than those obtained when looking at $c2o$ returns, and suggest significant intra-day continuation patterns.[9]

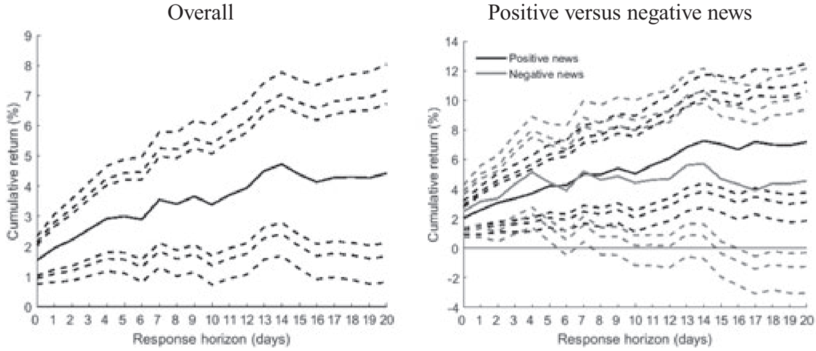
To investigate the degree to which our suggested news measure predicts asset prices beyond the day in which the news is published, we look at how news predicts cumulative close-to-close returns. In particular, let $y_{i,t:t+h}$ denote the cumulative close-to-close return for firm i across horizons t to $t + h$. Then, regression specification I from Table 1, which yielded the smallest short-term effect size, is estimated for each $h = 1, \ldots, 20$, using $y_{i,t:t+h}$ as the dependent variable. The left graph in Figure 2 reports the mean predictions together with 99, 95, and 90 percent confidence intervals from this experiment. By construction, the impact effect is as reported in Table 1, but the effect of a one unit news innovation also accumulates substantially over time. A maximum effect of roughly 4 percent is obtained after 15 business days, before it levels off. Converting this number into the effect following a one standard deviation news innovation gives an increase of roughly 7 basis points. Without exception, the response path is significant at the 1 percent level.[10]

Many textual studies in finance have given the news–return relationship a behavioral interpretation and documented significant overreaction patterns (Tetlock, 2014). In our results (see Figure 2), we see little sign of reversal, suggesting that the news topics carry new fundamental information (as

[9] In Table A5, in Online Appendix A, we run similar regressions for open-to-close returns ($o2c$), confirming that the intra-day effect of the news is roughly, depending on the exact model specification, between 50 and 130 basis points.

[10] We have also done these experiments by first cleansing the news topic variables for potential autocorrelation and common time fixed effects, giving the regressions an impulse response interpretation as in the local linear projection framework (Jordà, 2005). Doing so, we observe that the impact effect is of the same magnitude as already documented in Table 1, suggesting that the news topic variables are fairly exogenous to past developments in the market and not very persistent. Moreover, in the days following the initial news shock, the effect on returns accumulates as above, with little sign of reversal.

Figure 2. Predicted cumulative close-to-close returns



Notes: The left panel reports the results for overall news topic measure. The right panel reports the results when positive and negative news signals are allowed to have different effects. To facilitate the comparison with positive news, the response path following negative news is inverted. The solid line is the mean response to a one unit news innovation, and the dashed lines represent the 99, 95, and 90 percent confidence intervals, respectively. Standard errors are computed by clustering on trading day.

opposed to noise). An alternative explanation provided in the literature rests on an argument where heterogeneity among investors creates predictable patterns in the data, including overreaction to good news and underreaction to bad news (Frank and Sanati, 2018). However, when we allow for different slope parameters for positive and negative news innovations in the above regressions, we find no significant difference between the two response paths. As seen in the right panel in Figure 2, both positive and negative news innovations are associated with an initial underreaction, although this effect is somewhat weaker for negative news innovations.

One plausible interpretation of these results is given by theories of rational attention where information gathering is costly and/or the investors are cognitively constrained. In such a setting, the media matters because it can reach a broad population of investors and potentially alleviate informational frictions by contributing to information diffusion (Peress, 2014). Similarly, our results could be interpreted in line with a view where the media acts as an “information intermediary” between agents and the state of the world (Nimark and Pitschner, 2019), and editorial decisions play a more independent role. We explore further these subtle differences in interpretation in Section 4.

3.2. Randomization, additional fixed effects, and interaction terms

A novelty of our analysis is that we treat every day as a news day by linking news topics to returns using word distributions derived from the business newspaper and the firm's textual descriptions (see Section 2.3). Panel A of Table A3, in Online Appendix A, shows that the way we link companies to news is crucial for obtaining significant results. In particular, when we randomly assign news topics to firms, and run exactly the same regressions as described for Table 1, we find that almost no significant predictive power can be established.[11]

One might suspect, however, that the way in which we link firms to news resembles some type of industry classification, and that the results presented thus far capture industry effects (see, e.g., Hou and Robinson, 2006), or that the news topic variables proxy well-known weekday effects (see, e.g., Doyle and Chen, 2009). In Panel B, in Table A3, we redo the regressions from Table 1, but now include industry-specific dummies and control for the day of the week. As seen from the results, irrespective of which control variables we include, the news topic coefficients are almost identical to those found earlier.

Another concern could be that the news topic variables are associated with particular firm characteristics such as book-to-market value, size, or liquidity, which are well-known pricing factors (Fama and French, 1993; Carhart, 1997; Pastor and Stambaugh, 2003). In Table A4, in Online Appendix A, we interact the news variable with book-to-market values ($B/M_{t-1} : Topic_t$), market values ($MV_{t-1} : Topic_t$), and turnover ($Turn_{t-1} : Topic_t$), and include these as additional control variables in the panel regressions. As seen from the results, none of the interaction terms is significant, and the coefficients associated with the news topic variable remain significant at the 1 or 5 percent level with roughly the same effect size as already presented. Thus, there are no significant patterns indicating that our main results are driven by value, size, or liquidity characteristics.

We have also tried sorting firms into quantiles based on their average book-to-market value, size, and turnover. Then, for each quantile and firm characteristic, we estimated the effect of news. The findings resemble those described above, namely that for most quantiles and characteristics, the news coefficient is positive, significant, and shows no pattern of being associated with specific firm characteristics.

In sum, we find that our results are robust to falsification tests (randomizing topic assignments) and various additional fixed effects (industry and weekday effects), and are not driven by well-known firm characteristics.

[11] For close-to-close returns and the regression specification in Columns III of Table A3, the $Topic_t^R$ variable is barely significant at the 10 percent level.

Table 2. Firm-specific news topics and day t close-to-open and close-to-close returns across subsamples

	Close-to-open returns			Close-to-close returns		
	1996–2002	2002–2008	2008–2014	1996–2002	2002–2008	2008–2014
$Topic_t$	0.0169***	0.0084***	0.0086**	0.0267***	0.0135***	0.0099*
	(0.0047)	(0.0031)	(0.0039)	(0.0065)	(0.0042)	(0.0055)
R^2	0.0056	0.0100	0.0107	0.0045	0.0069	0.0062
Obs.	130,332	197,231	213,412	130,332	197,231	213,412
α_i	Yes	Yes	Yes	Yes	Yes	Yes
δ_t	Yes	Yes	Yes	Yes	Yes	Yes

Notes: For each return variable, the regression specification in Columns 1 and 4 from Table 1 is used. All regressions are estimated by OLS. Fixed effects are included as specified in the table. Following Tetlock et al. (2008), we compute clustered standard errors by trading day. Robust t-statistics are given in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

3.3. Changing market and media trends

During the last two decades, both the media and stock market have undergone substantial changes. First, as noted in Section 2.3, the breadth of the Norwegian stock market has become much bigger over the years, potentially suggesting that also a broader set of media is required to adequately cover it. For example, during the first years of our sample, fewer than 60 firms were listed on the Oslo Stock Exchange. In contrast, in 2014, over 120 companies were listed and included in our dataset. Second, while printed news was a primary media channel a decade ago, Internet usage and online consumption of newspaper content dominates today (Statistics Norway SSB, 2016). In terms of the number of readers of printed news, our primary source DN has been ranked as the fourth largest in Norway, irrespective of subject matter, throughout the whole sample. In terms of online readers, however, DN has faced substantially tougher competition. For example, DN's share of the total number of online readers declined by 25 percent around 2008 due to the establishment of competing news media.[12] Together, these trends suggest that DN's role as an information diffusion channel might be weakened across time, and the relationship between the (DN) news topics and returns accordingly.

The results reported in Table 2 address this issue. Here we have divided the sample (1996–2014) into three equally sized subsamples, and redone the estimation from the columns labeled I in Table 1. As seen from the table, the predictive effects are positive and significant for all subsamples and for both close-to-open and close-to-close returns, but the strength of

[12] See medianorway, Facts and figures on Norwegian media, https://www.medienorge.uib.no/english/.

the effect tends to diminish over time. The left panel in Figure 3 shows that this general pattern carries through also for longer-term predictions. During the period 1996–2002, we find a positive, highly persistent, and significant predictive relationship between news and returns. For the period 2002–2008, this relationship weakens somewhat, but remains significant. The really dramatic change is for the last subsample, 2008–2014, where the predictive relationship between news topics and returns becomes insignificant after only one day.

In line with the discussion above, one interpretation of these findings is that DN, from which we derive our news signal, has become less important for understanding the news–return relationship. However, the results might also indicate that financial markets have become more efficient over time, and that information frictions that were present during the 1990s and early 2000s are no longer binding.

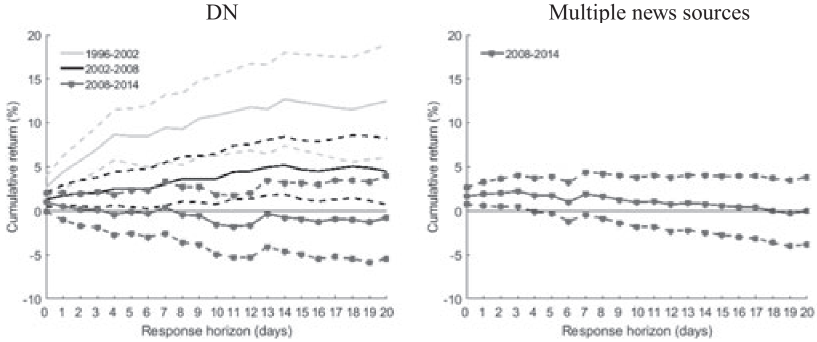
To cast further light on these two competing explanations, our textual data provider Retriever has provided us with a broad-based sample of news articles from the biggest players in the Norwegian (business) newspaper market. This extra set of data covers the period 2008–2014, and includes news from four additional sources.[13] We utilize these extra data in four steps. First, we clean the textual data, as described in Section 2.1. Then, in a second step, we apply a procedure for querying documents outside the set on which the LDA is estimated, as described in Section 2.3. That is, we keep the topic definitions estimated from the DN corpus, and classify the augmented corpus based on these existing word distributions. The advantage with this approach is that we ensure that the topics, in terms of word distributions, stay the same across the extended dataset (multiple sources) and the original one (DN only).[14] Third, we compute new topic time series, for the period 2008–2014, based on the tone and frequency associated with each topic from the aggregated corpus (DN and additional sources), as in Section 2.2. As such, the extra data allow us to capture a much broader news base than when using DN alone. Finally, we redo the predictive regressions discussed above.

The results reported in the right panel in Figure 3 are striking. When the broad-based news topic variables are used, they predict significant

[13]The sources are *Aftenposten*, *Finansavisen*, *Bergens Tidende*, and *E24*; the latter is an online media channel only. To avoid using news content that is generated as a response to market movements on day t, we define the online news corpus for a given day t as containing news articles from eight in the morning on day $t-1$ to eight in the morning on day t (i.e., before the market opens on day t).

[14]Because of lack of identifiability in the LDA, the estimates of the topic and word distributions cannot be combined across samples for an analysis that relies on the content of specific topics. A disadvantage of this approach is that by definition it does not take into account the possibility that the additional news sources write about other news topics than those defined by DN.

Figure 3. Predicted cumulative close-to-close returns across subsamples



Notes: The solid line is the mean response to a one unit news innovation, and the dashed lines represent the 95 percent confidence intervals, where standard errors are computed by clustered on trading day. In the left panel, DN is the only newspaper source. In the right panel, we include a broader set of news sources (available for the latest sample period only).

continuation, peaking after three business days and remaining significant for up to one business week. For longer horizons, the effect becomes insignificant, and slowly reverts to zero. This stands in stark contrast to the comparable result in the left panel in Figure 3, where the predictive relationship between news and returns during the 2008–2014 period was basically insignificant after day 0. Still, compared with a decade ago, the persistence of the news–return predictive relationship is reduced substantially.

We conclude from these results that the financial market might have become more efficient, but that the media still has significant predictive power. However, as the size of the stock market itself has grown, and the variety of news sources delivering business relevant news has proliferated, our results suggest that a broad-based news corpus now needs to be applied to capture informative news signals.

3.4. A news-based trading strategy

The analysis thus far has focused on average effects across all firms. To gauge the degree to which the news affects individual firms valuations, we run a zero-cost investment strategy similar to those implemented in, for example, Tetlock et al. (2008) and Boudoukh et al. (2013). The strategy is implemented as follows. For each trading day, we go US$1 long (in total) in all stocks that receive positive news in the morning, and US$1 short (in total) in all stocks that receive negative news. For a trade to take place, we require that we have at least five stocks on each side. Based on the continuation patterns shown in Figure 3, the stocks are held for up to five trading days. At the end of each day, we compute the total daily (close-to-close) return from both the long and short portfolios we have at that point in time, controlling for the fact that stocks are bought on opening prices and sold on close prices. The total daily return from the strategy is the difference between the daily return from the long and short portfolios.

Columns labeled I in Table A6, in Online Appendix A, summarize the yearly returns and Sharpe ratios generated by the benchmark zero-cost portfolio using DN and multiple media as news sources (from 2008), respectively. For the strategy utilizing only DN as a news source, negative returns are observed for five out of 18 years; 1997, 2006, 2007, 2009, and 2011. On average, across all the years, the annualized daily return is 16.8 percent, with a Sharpe ratio of about 0.89. For comparison, this return is almost four times that of the market as a whole; see Figure A3 in Online Appendix A, which has a Sharpe ratio of 0.33. Although good, these numbers improve substantially when the news signal traded upon utilizes multiple sources. In that case, negative portfolio returns are only observed in three out of 18 years, and the annualized average return is 29.1 percent with a Sharpe ratio of 1.56. However, as also seen from the table, the average numbers of daily trades conducted to form the long and short portfolios are substantial. In a real-world setting, this would have implied substantial trading costs, which would likely have subtracted away a large part of the aggregate returns.

To reduce the number of trades conducted, we also run an alternative trading strategy. This strategy is similar to that above, but with the difference that news is only traded upon if the news signal is over or below one standard deviation of the respective news topic time series. Here, the computations of the standard deviations are recursively updated throughout the trading experiment, using the past 252 observations to calculate the standard deviations. As seen from the columns labeled II in Table A6, this more restrictive trading strategy reduces the average returns on the portfolios somewhat. Still, the annualized average daily returns are 11.7 and 21.3 percent, with Sharpe ratios of 0.60 and 1.1, for the DN only and multiple sources strategies, respectively. More importantly, however, the alternative strategies generate these returns by far fewer trades than above.

Do the two zero-cost investment strategies generate risk-adjusted returns as well? In Table 3, we use the daily return series generated by the two strategies, subtract the risk-free rate, and run regressions controlling for the standard risk factors (Fama and French, 1993; Jegadeesh and Titman, 1993; Carhart, 1997; Pastor and Stambaugh, 2003): the market (MR), size

Table 3. Risk-adjusted return for zero-cost investment strategies

	DN		Multiple news sources	
	I	II	I	II
MR	−0.0445***	−0.0154	−0.0448***	−0.0110
	(0.0165)	(0.0149)	(0.0168)	(0.0156)
SMB	0.0557**	0.0060	0.0107	−0.0361
	(0.0250)	(0.0278)	(0.0257)	(0.0254)
HML	0.0584***	−0.0041	0.0516***	−0.0056
	(0.0194)	(0.0200)	(0.0189)	(0.0215)
UMD	0.0538***	0.0085	0.0373*	0.0037
	(0.0195)	(0.0188)	(0.0206)	(0.0205)
LIQ	0.1362***	0.0227	0.1312***	0.0420*
	(0.0217)	(0.0220)	(0.0225)	(0.0216)
$Alpha$	0.0333**	0.0267	0.0851***	0.0643***
	(0.0167)	(0.0178)	(0.0165)	(0.0178)
R^2	0.0601	0.0014	0.0409	0.0014
Obs.	4,667	4,667	4,688	4,688

Notes: Either DN only or multiple news sources are used to derive the news signal. Given the news source, in the columns labeled I, all news signals are potentially traded on. In the columns labeled II, only news signals over or above one standard deviation are traded on. The dependent variable is the strategy-generated return less the risk-free rate. The independent variables include contemporaneous factors for: the market (MR), size (SMB), book-to-market (HML), momentum (UMD), and liquidity (LIQ). We compute all coefficient standard errors using heteroskedasticity-consistent standard errors (White, 1980). Robust t-statistics are in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

(SMB), book-to-market (HML), momentum (UMD), and liquidity (LIQ).[15] Only for the alternative trading strategy, and when using DN as the only news source, can we not reject the null hypothesis of a zero $Alpha$. When using multiple news sources, for example, the point estimate is between 0.06 and 0.08, implying an annualized risk-adjusted daily return of between roughly 15 and 20 percent. Although comparisons across markets and time periods might be misleading, we note that these numbers are comparable in magnitude to those found in both Tetlock et al. (2008) and Boudoukh et al. (2013) for the US market. Interestingly, the news-based trading strategies tend to be negatively correlated with the market, and positively correlated with momentum and liquidity, but the significance of these correlations varies substantially.

We emphasize that the trading experiments conducted here are deliberately kept simple. More than providing examples of realistic trading

[15]Professor Bernt Arne Ødegaard, at Stavanger University, constructs these risk factors for the Norwegian market and makes them publicly available on the web site, Asset pricing data at OSE, https://ba-odegaard.no/financial_data/ose_asset_pricing_data/index.html (see also Ødegaard, 2017).

opportunities, they cast light on the robustness of the pooled time series regressions presented in the previous sections.[16] We conclude that the significant news–return relationship is not driven by the panel data approach, and also has the potential for being economically important.

4. The causal media effect

The news signal potentially contains (at least) two different components. First, news in the business newspaper can be genuine new information. Second, the media might itself affect markets by how they report news stories and by disseminating information to a broad population of investors. As genuine new information is more likely to be generated exogenous to the media (and reported in the media with a time lag), it is the second component that reflects media's potential causal role in predicting returns. To separate between these two components, however, is difficult, because we only observe the signal, and not its two underlying components.

To address this issue, we exploit a strike in the Norwegian newspaper market in 2002, which started on 30 May and ended on 7 June (i.e., lasting for seven business days). The same event was used in Peress (2014) to investigate the causal effect media has on trading and price formation. But, in contrast to his cross-country event study, we focus on the case of Norway and on changes in returns, and we condition our analysis on the news topic variables, and conduct a difference-in-difference identification strategy to isolate the potential negative consequences of the strike itself.[17] This allows us to obtain a novel estimate of the media effect in a given predictive relationship. Simply put, we ask how much of the increases in returns documented in the preceding sections can be attributed to the causal (DN) news topics media effect.

According to Peress (2014), the newspaper strike affected the press on a national scale, involved the media sector only, and occurred on days on which the stock market was open. Moreover, the strike was called by the media profession itself due to their working conditions, and it was not driven by stock market movements on the day of the strike or the preceding days. Thus, we can safely assume that it was truly exogenous to market developments.[18] Accordingly, we follow an event study approach where in total 103 stocks enter our sample in the year(s) prior to, during,

[16]As such, it is perhaps interesting to know which news is actually traded on across time. This is illustrated in Figure A4, in Online Appendix A, using a heatmap.

[17]That is, a strike might affect returns simply because strikes are bad economic events, and not because of the media blackout.

[18]See Peress (2014) for a richer discussion about these issues. It should be noted, however, that he also includes a Norwegian newspaper strike in 2004. During this journalist strike, the DN

Table 4. Summary statistics of news and returns, before and during and after the strike

	$E(X)$	$VAR(X)$	$Skew(X)$	$Kurtosis(X)$	$\min(X)$	$\max(X)$	N
Δr^{c2o}_{d-ba}	−0.3918	1.1684	0.0028	4.6416	−3.8829	3.0488	103
Δr^{c2c}_{d-ba}	−0.6171	1.5688	−0.2377	4.6911	−4.5117	3.9407	103
$t(\beta)$	2.3016	3.9414	0.6000	3.3606	−1.3632	8.8093	103

Notes: In the rows labeled Δr_{d-ba}, each table entry is a function of $\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba}$, where $\hat{r}$ is the average return for company i, during the strike period ($\hat{r}_{i,d}$), and before and after ($\hat{r}_{i,ba}$), respectively. The length of the strike period ($\hat{r}_{i,d}$) is constant and equal to seven business days. The total length of the window used during and before and after the strike equals 51 days. In the interest of readability, the numbers reported in the columns $E(X)$, $VAR(X)$, $min(X)$, and $max(X)$ are scaled by 100. The table row labeled $t(\beta)$ provides summary statistics for $t_i(\beta)$, a standardized firm-specific news loading, estimated on a classification sample prior to the strike period.

and after the strike. We focus on both their close-to-open and close-to-close returns, and use N days prior to and N days after the strike to compute the non-strike affected returns. In the following, we denote the change in returns $\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba}$, where $\hat{r}$ is the average return for company i, during the strike period ($\hat{r}_{i,d}$) and before and after ($\hat{r}_{i,ba}$), respectively. By adjusting N, we can down-weight observations right before and right after the strike, as these might be driven by anticipation effects (about the forthcoming strike) and adjustments following the end of the strike period. We denote the total length of the sample before, during, and after the strike by W. In the main results presented below $W = 51$, but we show that our results are robust to both shorter and longer windows.

4.1. Unconditional and conditional effects

The first row in Table 4 documents that, on average across firms, the average returns during the strike period fell by between 40 and 60 basis points relative to the average returns prior to and after the strike. However, the dispersion across firms is large, with minimum and maximum values reaching −3.88 and 3.04 percent, and −4.51 and 3.94 percent, for $c2o$ and $c2c$ returns, respectively. Still, the skewness and kurtosis statistics suggest that the distribution is not far from normal, albeit with some outliers. For both types of returns, the mean effects are significantly different from zero on the 1 percent significance level, suggesting that the media has a positive causal role in explaining short-term return patterns.[19]

newspaper was in fact published, and the event cannot be used here. We further note that, given the timing of the strike event in 2002, the explosion of digital media seen the last decade had hardly begun.

[19]A direct comparison of our results to those in Peress (2014) would have been interesting, but not feasible. He focuses on pooled cross-country averages, trading volume, and volatility, and does not report raw return statistics for the Norwegian newspaper strike in 2002.

A valid objection to the simple calculations done above is that they do not necessarily tell us something about how the media shortage affected returns. The fall might simply be due to the effect that the strike itself, or other not controlled for common events, had on markets. Accordingly, we would ideally need two different groups of firms and returns: one where the media shortage should matter, and one where it should not. Unfortunately, such classifications are not observed. Moreover, although the numbers might reflect the causal media effect, they do not necessarily relate to the news topic variables used in this study (i.e., those obtained from the DN newspaper).

To accommodate the concern, and to relate the media shortage to the DN news topic variables, we run a difference-in-difference type of experiment, with some modifications. As above, we first compute the difference between returns during the strike and those before and after the strike. Then, conditioning on how sensitive the respective stocks were to the news topic variables in the year prior to the strike, we construct a treatment and control group, and run simple regressions to quantify the media effect due to the shortfall of the DN news topics. Intuitively, all stocks might be affected by the strike, but those firms that had a particularly high sensitivity to news topics prior to the strike should also respond strongly to their shortfall during the strike.[20] More formally, we consider the model,

$$\hat{r}_{i,e} = \alpha_i + \delta D_e + \tau w_{i,e} + u_{i,e}, \tag{2}$$

where the event indicator $e = \{ba, d\}$ indicates the periods before and after (ba) and during (d) the strike, α_i is a firm-fixed effect constant across e, $D_e = 1$ if $e = d$ and zero otherwise, and $u_{i,e}$ are idiosyncratic errors. The parameter of interest is τ, measuring the effect of $w_{i,e}$, a binary indicator of the treatment. Before and after the strike $w_{i,ba} = 0$ for all i. During the strike, however, $w_{i,d} = 1$ if firm i is in the treatment group (i.e., particularly sensitive to the DN news topics), and zero otherwise. A simple estimation procedure of the two-period model in equation (2) is to first difference to remove α_i,

$$\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba} = \delta + \tau \Delta w_i + \Delta u_i, \tag{3}$$

with $\Delta w_i = w_{i,d}$ (as $w_{i,ba} = 0$ for all firms i in period $e = ba$). Conditional on the strike being truly exogenous, the central premises for this experiment are: (i) that news in terms of new information was released also during the strike period, although not through the mass media; (ii) that the distribution

[20] Estimating the average media shortage effect only for returns in the treatment group would not permit us to exclude the general effect the strike itself might have on returns. Of course, if the effect of the general strike affects firms in the two groups differently, our experiment design will not be able to efficiently isolate the strike effect from the (DN) media shortage effect.

of new information, in terms of topics and tonality, is close to that observed in non-strike periods. As such, a rejection of the null hypothesis of no significant differences in returns between the firms in the control and treatment groups (i.e., a significant τ estimate) imply a causal media effect.

While equation (3) is a standard difference-in-difference model (Meyer, 1995; Angrist and Krueger, 1999), the crux here is to construct $w_{i,d}$, which depends on the firm's news topic sensitivity. We construct $w_{i,d}$ in two steps. First, we estimate news topic sensitivity, denoted by $t_i(\beta)$, from time series regressions of each individual firm's close-to-open return ($y_{i,t}$) on news topics ($T_{k,t}$) during the year preceding the strike,

$$y_{i,t} = \beta_i T_{k,t} + z'_{i,t}\lambda + u_{i,t}, \tag{4}$$

where $t_i(\beta)$ is the t-statistic associated with β_i.[21] The third row of Table 4 reports how $t(\beta)$ is distributed across the 103 firms in the sample. Clearly, the t-statistic is large (and significant) on average. Still, many companies also have a negative exposure towards the news variable, although not significantly so. Second, given the distribution of $t(\beta)$, we define a cut-off co, and set $w_{i,d} = 1$ if $t_i(\beta) > co$ and $w_{i,d} = 0$ otherwise. Here, the natural cut-off is $co = 2$ (i.e., approximately the 5 percent significance level), dividing the sample in roughly two equally sized groups. We show below, however, that our results are robust to a range of other plausible cut-off values (more or less significant).

A priori it is reasonable to assume that large firms, and firms with a high turnover, have a stronger news exposure. These priors are also reflected in our results. As seen in Figure A5a, in Online Appendix A, there is a clear positive relationship between $t(\beta)$ and firm size and turnover. However, given the co cut-off, we do not find any systematic differences between firms with $t_i(\beta) > co$ or $t_i(\beta) <= co$ and their news topic mapping. This is illustrated in Figures A5b and A5c. That is, firms in both groups are mapped to more or less the same type of most news topics. Still, there are some differences in the relative weights, and firms with $t_i(\beta) <= co$ tend to relate to a more diverse set of news topics.

Column I in Table 5 reports the results from estimating equation (3) using $w_{i,d}$ and two different dependent variables, $\Delta r^{c2o}_{i,d-ba}$ and $\Delta r^{c2c}_{i,d-ba}$. For close-to-open returns, we find no significant differences in means across the two groups of firms. For close-to-close returns however, there is a clear and significant difference. Firms where the news topic variables were important prior to the strike period experience 57 basis points lower returns during

[21] We focus on $t(\beta)$, rather than β, to control for differences in precision as a result of differences in residual variance. To reduce potential biases in the regressions, we also include additional controls, $z_{i,t}$, including R^{mi}_{t-1}, $B/M_{i,t-1}$, $MV_{i,t-1}$, and lagged close-to-close returns for stock i (i.e., the significant regressors in the panel regressions run in Section 3).

Table 5. Estimated media effects from $\Delta r_{i,d-ba} = \delta + \tau \Delta w_i + \Delta u_i$

Return	I			Size	Power
	δ	τ	R^2adj	τ	τ
$\Delta r^{c2o}_{i,d-ba}$	-0.0026^{*}	-0.0027	0.0055	0.0000	0.3478
	[0.076]	[0.242]			
$\Delta r^{c2c}_{i,d-ba}$	-0.0033^{**}	-0.0057^{**}	0.0423	0.0870	0.5217
	[0.020]	[0.034]			

Notes: Here, $\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba}$, and $\hat{r}$ is the average return for company i, during the strike period ($\hat{r}_{i,d}$) and before and after ($\hat{r}_{i,ba}$), respectively. The total window length is $W = 51$. Δw_i is a binary variable with $\Delta w_i = 1$ if $t_i(\beta) > co$, and zero otherwise. $co = 2$, and $t_i(\beta)$ is a standardized firm-specific news loading, estimated on a classification sample prior to the strike period. The Size and Power columns report the fraction of test statistics with a p-value <0.05 in a simulation experiment on time periods without an actual strike. See the text for details. The p-values, reported within the square brackets, are computed using a residual bootstrap taking into account estimation uncertainty in $t_i(\beta)$ and potential heteroskedasticity (White, 1980) in the second-stage regressions. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

the media shortage relative to those firms where the news topic variables were not important.

Qualitatively, none of the results reported in columns I in Table 5 are affected by varying the cut-off value (co) between 1.6 and 3 (from weakly significant to highly significant) when constructing $w_{i,d}$, or by using a different window length to construct $\hat{r}_{i,ba}$ (see Tables A7 and A8, in Online Appendix A). Similarly, controlling for a stock's overall market exposure ($Beta$) and the (average) characteristics B/M, MV, and $Turn$ when estimating equation (3) results in lower adjusted R^2 and reduced significance level of the estimated coefficients, but does not change our qualitative conclusions regarding τ and close-to-close returns. For example, in the augmented regression, $\tau = -0.0043$, compared to -0.0057 in Table 5, and the p-value is 0.086 (see Tables A9 and A10).

Moreover, it seems very unlikely that the patterns documented above are obtained by chance, reflect differences in trends, or are present regardless of the media shortage. We show this by sampling 23 non-overlapping periods of returns with $W = 51$, computing $\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba}$ as if $r_{i,d}$ contained a strike event (while, in reality, it did not), and redo regression (3) for each draw.[22] On average, when no actual strike is present, we are only able to reject the null-hypothesis of no significant effects at the 5 percent level in at most 8.7 percent of the cases (see column "Size" of Table 5).

[22]The first period drawn starts in early 2000, and the last period drawn ends in late 2005. The window with the actual strike period in 2002 is excluded. Increasing the number of non-overlapping periods, by drawing non-overlapping periods from a larger time-span, means that many firms that were included in the sample in 2002 will fall out (they are either not listed on the Oslo Stock Exchange, or delisted). In unreported results, we show that the size and power statistics are robust to using a shorter window (i.e., with $W = 31$) and a larger number of non-overlapping periods.

Conversely, if we impose the media effect estimated above on each of the $\hat{r}_{i,d}$ periods sampled, we see from the column labeled "Power" that we obtain a significant relationship in 52 percent of the cases ($c2c$), and a substantially lower number for the effect that was not significant in the first place ($c2o$). Naturally, if we increase the effect size imposed on the non-overlapping periods, the power increases, and vice versa (not shown).

Together, these results provide strong evidence towards a causal (DN-specific) news topic media effect. For close-to-close returns, the difference in mean between the synthetic control and treatment groups is even of the same magnitude as the total strike effect documented in Table 4 (i.e., 60 basis points). To put these numbers into context, back-of-the-envelope calculations suggest that up to 20 percent of the close-to-close returns predicted by news topics are due to the media effect alone.[23]

4.2. Accounting for treatment intensity and asymmetries

A weakness with the approach just taken is that it does not account for the basic intuition that, all else equal, firms' news topic sensitivity might affect the intensity at which the media shortage affects returns. For example, firms with a particularly high (low) and significant (insignificant) sensitivity to news prior to the strike period might also be more negatively (positively) affected than the other firms during the media shortage.

To investigate this hypothesis, we extend the regression model in equation (2) by taking into account potential differences in how the media shortage affects returns within and between the treatment and control group. In particular, we consider the model,

$$\hat{r}_{i,e} = \alpha_i + \delta D_e + \tau^t \tilde{w}_{i,e} t_i(\beta)^t + \tau^c \tilde{w}_{i,e} t_i(\beta)^c + \tau \tilde{w}_{i,e} + b_2 t_i(\beta)^t + b_3 t_i(\beta)^c + u_{i,e}, \tag{5}$$

where α_i, δD_e, and $u_{i,e}$ have the same interpretations as before. Now, however, $\tilde{w}_{i,e}$ is a binary strike indicator, where $\tilde{w}_{i,e} = 1$ if $e = d$ and zero otherwise for all firms i. This strike event indicator is then interacted with the terms $t_i(\beta)^t$ and $t_i(\beta)^c$. Also, $t_i(\beta)^t = t_i(\beta)$ if $t_i(\beta) > co$ and zero otherwise, and $t_i(\beta)^c = t_i(\beta)$ if $t_i(\beta) <= co$ and zero otherwise. As such, in response to the media shortage, τ^t and τ^c capture the potential asymmetries between firms with a significant and non-significant sensitivity to the news topics. For estimation, we first difference equation (5), yielding the following model,

[23] According to the estimates in Table 5, the DN news topic effect is 57 basis points, while the minimum predictive effect from the pooled time series regression reported earlier in Table 1 is 153 basis points (i.e., $57/153 \approx 0.4$). However, the standard deviation of weekly news topics (i.e., the duration of the strike) is between two and three times as large as at the daily frequency, which implies an upper estimate around $0.4/2$.

Table 6. Estimated media effects from $\Delta\hat{r}_{i,d-ba} = \tilde{\delta} + \tau^t t_i(\beta)^t + \tau^c t_i(\beta)^c + \Delta u_i$

Return	I				Size		Power	
	$\tilde{\delta}$	τ^t	τ^c	$R^2 adj$	τ^t	τ^c	τ^t	τ^c
$\Delta r^{c2o}_{i,d-ba}$	−0.0041**	−0.0004	0.0028	0.0372	0.1304	0.0435	0.1304	0.4348
	[0.046]	[0.412]	[0.434]					
$\Delta r^{c2c}_{i,d-ba}$	−0.0026	−0.0018**	−0.0003	0.0752	0.435	0.0435	0.5652	0.0435
	[0.258]	[0.024]	[0.904]					

Notes: Here, $\Delta r_{i,d-ba} = \hat{r}_{i,d} - \hat{r}_{i,ba}$, and $\hat{r}$ is the average return for company i, during the strike period ($\hat{r}_{i,d}$) and before and after ($\hat{r}_{i,ba}$), respectively. The total window length is $W = 51$. Also, $t_i(\beta)$ is a standardized firm-specific news loading, estimated on a classification sample prior to the strike period. $t_i(\beta)^t = t_i(\beta)$ if $t_i(\beta) > co$ and zero otherwise. $t_i(\beta)^c = t_i(\beta)$ if $t_i(\beta) <= co$ and zero otherwise. In both cases, $co = 2$. The Size and Power columns report the fraction of test statistics with a p-value<0.05 in a simulation experiment on time periods without an actual strike. See the text for details. The p-values, reported within the square brackets, are computed using a residual bootstrap taking into account estimation uncertainty in $t_i(\beta)$ and potential heteroskedasticity (White, 1980) in the second-stage regressions. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

$$\Delta\hat{r}_{i,d-ba} = \tilde{\delta} + \tau^t t_i(\beta)^t + \tau^c t_i(\beta)^c + \Delta u_i, \qquad (6)$$

where $\tilde{\delta} = \delta + \tau$, and apply ordinary least squares.

Columns I of Table 6 report the regression output. Starting with the results in the $c2c$ row, we see that the $t_i(\beta)^t$ term is highly significant and, based on the argument above, has the expected negative sign. However, close-to-close returns for firms in the control group are unaffected by the media shortage. As such, these results are consistent with the results presented in the previous section. Moreover, as the average t-statistic in Table 4 is well above 2, the effect of the media shortage, as a percentage of the total predictive effect, is approximately the same as previously reported. Interestingly, this pattern is reversed when we look at close-to-open returns, where firms in the treatment group experience an extra underreaction, while the firms in the control group experience an initial overreaction and subsequent intra-day reversal. However, these effects are very uncertain and far from significant. Size tests, conducted as described in the previous section, but now applied to the model in equation (6), show that these results are highly unlikely to occur in time periods without a strike (see the column labeled "Size" in Table 6).

In sum, the results presented here and in Section 4.1 give media an important causal role for understanding asset price fluctuations. Positive evidence of the media's causal role in financial markets has also been documented in Engelberg and Parsons (2011), Dougal et al. (2012), Peress (2014), and Li (2018). Engelberg and Parsons (2011) analyze trading volumes, and show that trades by individual investors located in different locations in the US respond to local newspaper coverage. Dougal et al. (2012) use exogenous variation in the identity of *Wall Street Journal* columnists, and show that this is a good predictor of the next-day return on the Dow Jones Industrial Average. While they appeal to a sentiment story whereby the bullish or bearish sentiment conveyed by columnists influence investors, our results provide evidence more in line with an informational dissemination explanation where the media is an important channel for broadcasting fundamental information. As such, our findings are more consistent with the studies by Peress (2014) and Li (2018), who find evidence showing, respectively, that the media influences the stock market by increasing the speed with which information diffuses across investors and that professional investors increase their information production and potential profits when media coverage increases. We differentiate ourselves by potentially giving the media a larger independent role in the news-return relationship. In particular, during the strike, a media information shortfall seems to result in an adverse effect relative to the counterfactual.

5. Conclusion

News in business newspapers predicts daily returns, and the media has an important causal role in this predictive relationship. We reach these conclusions after decomposing the corpus in the main Norwegian business newspaper into daily news topics and linking them to firms and returns.

Although the news topics are available in the morning, well before the market opens, we document significant underreaction in market prices to news and clear patterns of continuation in the days following the initial news release. These results hold both in pooled time series regression, simple zero-cost news-based trading strategies, and are robust when controlling for numerous commonly used predictors.

Further, by exploiting an exogenous strike in the Norwegian newspaper market, in 2002, we are able to isolate the media component of the news signal from the new information component. Returns for individual firms with a significant exposure to our news measures fall by 57 basis points during the strike period relative to firms with an insignificant news topic exposure. Because the average firm in the sample has a positive exposure to news, our estimates suggest that the news media component plays a significant role, and that up to 20 percent of the predictive effect is due to the causal media effect.

In sum, our analysis supports a view where the media acts as an “information intermediary” between agents and the state of the world, and where editorial decisions potentially play an independent role in the news-return relationship. As such, our analysis speaks to a small, but fast growing, body of literature in economics using natural language processing technologies to establish predictive and causal relationships between news media and economic outcomes. As the methodology applied here is general, exciting avenues for future research include expanding the scope of the analysis to other countries, and investigating in greater detail how changing media trends might affect the relationship between news media and financial markets.

Supporting information

Additional supporting information can be found online in the supporting information section at the end of the article.

Online appendices
Replication files

References

Angrist, J. D. and Krueger, A. B. (1999), Empirical strategies in labor economics, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 3, Chapter 23, Elsevier, Amsterdam, 1277–1366.

Antweiler, W. and Frank, M. Z. (2006), Do US stock markets typically overreact to corporate news stories?, available at SSRN, https://ssrn.com/abstract=878091.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003), Latent Dirichlet allocation, *Journal of Machine Learning Research 3*, 993–1022.

Boudoukh, J., Feldman, R., Kogan, S., and Richardson, M. (2013), Which news moves stock prices? A textual analysis, National Bureau of Economic Research (NBER) Working Paper 18725.

Calomiris, C. W. and Mamaysky, H. (2017), How news and its context drive risk and returns around the world, Research Paper 17-40, Columbia Business School.

Campbell, J. Y., Grossman, S. J., and Wang, J. (1993), Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics 108*, 905–939.

Carhart, M. M. (1997), On persistence in mutual fund performance, *Journal of Finance 52*, 57–82.

Chang, J., Gerrish, S., Wang, C., Boyd-graber, J. L., and Blei, D. M. (2009), Reading tea leaves: how humans interpret topic models, in Y. Bengio et al. (eds), *Advances in Neural Information Processing Systems 22*, MIT Press, Cambridge, MA, 288–296.

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

Doyle, J. R. and Chen, C. H. (2009), The wandering weekday effect in major stock markets, *Journal of Banking & Finance 33*, 1388–1399.

Engelberg, J. E. and Parsons, C. A. (2011), The causal impact of media in financial markets, *Journal of Finance 66*, 67–97.

Fama, E. F. and French, K. R. (1993), Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics 33*, 3–56.

Frank, M. Z. and Sanati, A. (2018), How does the stock market absorb shocks?, *Journal of Financial Economics 129*, 136–153.

French, K. R. and Roll, R. (1986), Stock return variances, *Journal of Financial Economics 17*, 5–26.

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

Hansen, S., McMahon, M., and Prat, A. (2018), Transparency and deliberation within the FOMC: a computational linguistics approach, *Quarterly Journal of Economics 133*, 801–870.

Heinrich, G. (2009), Parameter estimation for text analysis, Technical report, Fraunhofer IGD.

Hou, K. and Robinson, D. T. (2006), Industry concentration and average stock returns, *Journal of Finance 61*, 1927–1956.

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

Jordà, Ò. (2005), Estimation and inference of impulse responses by local projections, *American Economic Review 95* (1), 161–182.

Kacperczyk, M., Van Nieuwerburgh, S., and Veldkamp, L. (2009), Rational attention allocation over the business cycle, National Bureau of Economic Research (NBER) Working Paper 15450.

Larsen, V. H. (2021), Components of uncertainty, *International Economic Review 62*, 769–788.

Larsen, V. H. and Thorsrud, L. A. (2019), The value of news for economic developments, *Journal of Econometrics 210*, 203–218.

Larsen, V. H., Thorsrud, L. A., and Zhulanova, J. (2021), News-driven inflation expectations and information rigidities, *Journal of Monetary Economics 117*, 507–520.

Li, B. Y. (2018), Beyond attention: the causal effect of media on information production, working paper available at SSRN https://ssrn.com/abstract=3200815.

Loughran, T. and Mcdonald, B. (2011), When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance 66*, 35–65.

Meyer, B. D. (1995), Natural and quasi-experiments in economics, *Journal of Business & Economic Statistics 13*, 151–161.

Næs, R., Skjeltorp, J. A., and Ødegaard, B. A. (2009), What factors affect the Oslo Stock Exchange?, Working Paper 2009/24, Norges Bank.

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

Ødegaard, B. A. (2017), Empirics of the Oslo Stock Exchange: asset pricing results 1980–2016, UiS Working Papers in Economics and Finance 2017/2, University of Stavanger.

Pang, B., Lee, L., and Vaithyanathan, S. (2002), Thumbs up? Sentiment classification using machine learning techniques, in *Proceedings of the Association for Computational Linguistics (ACL) 2002 Conference on Empirical Methods in Natural Language Processing, EMNLP '02*, Volume 10, ACL, Stroudsburg PA, 79–86.

Pastor, L. and Stambaugh, R. F. (2003), Liquidity risk and expected stock returns, *Journal of Political Economy 111*, 642–685.

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

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

Petersen, M. A. (2009), Estimating standard errors in finance panel data sets: comparing approaches, *Review of Financial Studies 22*, 435–480.

Roll, R. (1988), R^2, *Journal of Finance 43*, 541–566.

Schmidt, D. (2013), Investors' attention and stock covariation: evidence from Google sport searches, Working paper, HEC Paris.

Shiller, R. J. (2017), Narrative economics, *American Economic Review 107* (4), 967–1004.

Statistics Norway SSB (2016), Norwegian media barometer, 2015, http://www.ssb.no/en/kultur-og-fritid/statistikker/medie/aar/2016-04-14#content.

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

Tetlock, P. C. (2014), Information transmission in finance, *Annual Review of Financial Economics 6*, 365–384.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008), More Than words: quantifying language to measure firms' fundamentals, *Journal of Finance 63*, 1437–1467.
Thorsrud, L. A. (2020), Words are the new numbers: a newsy coincident index of the business cycle, *Journal of Business & Economic Statistics 38*, 393–409.
White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica 48*, 817–838.

First version submitted May 2019;
final version received November 2021.