What do we Learn from a Machine Understanding News Content? Stock Market Reaction to News*

Marie Brière[†] Karen Huynh[‡] Olav Laudy[§] Sébastien Pouget[¶]

October 19, 2022

Abstract

Using textual data extracted by Causality Link platform from a large variety of news sources (news stories, call transcripts, broker research, etc.), we build aggregate news signals that take into account the tone, the tense and the prominence of various news statements about a given firm. We test the informational content of these signals and examine how news is incorporated into stock prices. Our sample covers 1,701,789 news-based signals that were built on 4,460 US stocks over the period January 2014 to December 2021. We document large and significant market reactions around the publication of news, with some evidence of return predictability at short horizons. News about the future drives much larger reactions than news about the present or the past. Stock returns also react more to high-coverage news, fresh news and purely financial news. Finally, firms' size matters: stocks that are not components of the Russell 1000 index experience larger reactions to news compared to those that are Russell 1000 components. Implications of our results for financial analysts and investors are offered and related to the links between news, firms' market value and investment strategies.

Keywords: Natural Language Processing, Textual Analysis, Efficient Market Hypothesis, ESG

---

*We would like to thank Elise Boutin for collaborating on a preliminary version of this paper. We are also grateful to Pasquale Galassi, Matteo Germano, Pierre Haren, Eric Jensen, Matthieu Keip, Rami Mery, and Sofia Sakout for useful comments. Sébastien Pouget gratefully acknowledges the financial support of Causality Link in partnership with TSE-P, a research foundation under the aegis of the Toulouse School of Economics, and of the Agence Nationale de la Recherche (ANR-16-CE26-0008-01, project PIMS).

[†]Amundi Asset Management, Paris-Dauphine University, PSL Research University, Université Libre de Bruxelles. E-mail: marie.briere@amundi.com

[‡]Amundi Asset Management. E-mail: karine.huynh-ext@amundi.com

[§]Causality Link. E-mail: olav.laudy@causalitylink.com

[¶]Toulouse School of Economics and Toulouse School of Management, University of Toulouse Capitole. E-mail: sebastien.pouget@tse-fr.eu

1 Introduction

Advances in natural language processing (NLP) algorithms have made it increasingly easy to extract relevant content from a large volume of textual data available in the market. Quantifying news content that is relevant with NLP has a great advantage compared to traditional event studies analysis. It allows to test the impact of a large variety of events, not only specific news announcement such as macroeconomic release, earnings, mergers and acquisitions, dividend changes or analysts' recommendations. News announcement surprisingly have only a limited impact on stock prices (Change and Change (1983),Cutler et al. (1988), Campbell (1990)). One reason is that firm fundamentals can also be affected by non quantifiable information that is not necessarily reflected in accounting data, for example managers' communication, or firm's ESG policy. Finally, measuring news content also allows to detect variations in market sentiment that is not related to the arrival of a particular news.

Several studies have already shown that signals (or scores) constructed using textual analysis techniques are informative and predictive of stock returns (For example, Tetlock (2007), Tetlock et al. (2008), Loughran and McDonald (2011)). These studies mainly focus on examining the quantitative content from a given source of information at a time. Therefore, their extracted signals mainly reflect the positiveness or negativity of the tone used in each document.

In this paper, we use the Artificial Intelligence (AI) platform of Causality Link used to exploit firms' fundamental information from news on a much larger scale. Specifically, for a given company and a given day, the system can collect textual data from different sources including news stories, call transcripts, broker research, etc. This enables us to build an aggregate news signal that captures not only the positive or negative tone of news but also how popular such news is in the market that day. In this study, we first aim to explore the informational content of the aggregate news signal. Secondly, we investigate how and when new fundamental information is incorporated into prices. Thirdly, we explore the possible heterogeneity of such price reactions over various dimensions.

For every day and for each listed firm in the US from 2014 to 2021, we construct an aggregate news signal based on the positiveness or negativity of each news statement and the number of statements associated with that firm. Then, for each trading day, we sort all stocks according to their signal. We construct a long portfolio consisting of stocks whose aggregate signals are above the 90^{th} percentile (firms with good news) and a short portfolio consisting of stocks whose aggregate signals are below the 10^{th} percentile (firms with bad news). Looking at the returns on such portfolios around the news publication days allows us to test the informational content of the aggregate signal and measure how news is incorporated into prices. We apply the same procedure to different subsets of data based on news coverage, news novelty, types of news (financial and ESG news), tenses of news (past, present and future) and firms' sizes.

Our results are as follows. *First*, we show that our aggregate news signal can capture the content of news. Specifically, on the day when news related to a particular firm is published (day 0), we observe an average portfolio daily return of 0.73% (significant) for the long portfolio and of -0.54% (significant) for the short portfolio. This corresponds to an abnormal return (i.e., Fama-French three-factor alpha) of 1.27% for the long-short strategy.

Second, our results provide some evidence of market reaction before the news is published. The long portfolio experiences a daily return of 0.41% (significant) and the short portfolio a return of -0.34% (significant) on day -1. Fresh news, on the other hands, receives much smaller market reaction on day -1 compared to stale news. This observation suggests that the delay in publications of some news stories plays a role in generating early market reaction together with investors' anticipation of the news, perhaps thanks to private information.

Third, we show that it is possible to benefit from a trading strategy based on our signals depending on the nature of news and firm size. For example, one may earn profit from executing a long-short strategy on the sub-sample of fresh news, financial news or news about mid-cap stocks. Although the result suggests that it is possible to build a profitable strategy on the aggregate news signals, one needs to check if transaction costs would not erase these abnormal returns.

Fourth, we document market reactions to news about the past, the present and the future around day 0. In general, the magnitudes of such market reactions are the largest for news about the future: 1.69% for average returns on day 0 compared to 1.15% and 0.43% for news about the present and the past, respectively. This result is consistent with our expectation that news about the past or the present is more likely to be already incorporated into stock prices despite not having made it to the media and news sources that we consider. Additionally, news about the near future triggers a market reaction that is about three times larger compared to news about the distant future. We verify that it is unlikely that the types of news associated with short-term and long-term horizons drive the difference in market reaction. Our interpretation is that news about the near future contains less uncertainty or that investors discount the change in cash flows less heavily for this type of news.

Fifth, we show that news with higher coverage from the media receives monotonically more market reaction compared to those with lower coverage, controlling for firm size. Specifically, the long-short portfolio of Russell 1000 stocks on very high-coverage news (more than 10 mentions of the same Key Performance Indicator, KPI) generates a day-0 average return of 2.03% (significant), almost ten times larger compared to the one on very low-coverage news (with only one mention of the same KPI) at 0.22% (significant). This result is consistent with the expectation that important news which receives more public attention should receive a stronger market reaction compared to insignificant one.

Sixth, fresh news is shown to generate much stronger market reaction on the release day (day 0) compared to stale news. Specifically, the average returns on the long-short portfolio on day 0 are 1.77% (significant) for fresh news and 1.20% (significant) for stale news. These findings are consistent with our expectation that fresh news should deliver more surprises to the market on the release day compared to stale news. As expected, there are much less early market reaction for fresh news at day t-1 (0.32%, significant) compared to stale news (0.81%, significant).

Seventh, we find weak positive (yet significant) reactions around day 0 to both good and bad news regarding environmental, social and corporate governance (ESG) issues. These surprising market reactions may be driven by investors' preference for firms' disclosure of ESG information, being good or bad. Additionally, it may also be driven by good financial news that is released in the same day.

Eighth, we find that the magnitude of market reaction to news also depends on firm size. Specifically, we show that a long-short strategy based on the universe of firms with large market capitalization generates positive and significant returns on day 0 (0.67%) but of much smaller magnitude compared to the rest of the non-microcaps firms (1.79%). Information regarding smaller firms is more opaque to the market as these firms tend to have less analyst coverage. This could explain why news has a larger impact on prices. Another possible explanation is that, for small firms, outstanding news are likely to be the only news that are made available and hence one should observe stronger reactions to such news.

Our paper contributes to the rich literature on the informational content of news signals extracted from different textual sources, either as a proxy for market sentiment or firm-specific fundamental information. In a pioneer study on the relation between market sentiment and stock returns, Tetlock (2007) relies on the word counts of the Wall Street Journal columns grouped according to a pre-specified sentiment dictionary (e.g: the General Inquirer's Harvard IV-4) to build a measure of market pessimism. He documents a "two-way" relationship between market returns and pessimistic sentiment. Specifically, pessimism is associated with low future returns but pessimism itself also follows negative returns. Interestingly, there is evidence of a reversal in returns after pessimistic sentiment, which implies that downward price pressure is temporary driven by investors' sentiment rather than by new or stale information that has not yet been incorporated into prices.

Signals extracted from firm-specific fundamental news have also been shown to be informative and able to predict returns following news publication. Tetlock et al. (2008) uses the fraction of negative words in news associated with a particular firm as a measure for the news negativity. They show that a higher fraction of negative words is associated with more negative stock returns on the following trading day. Their long-short strategy based on the fraction of negative words produces positive and significant abnormal returns. Loughran and McDonald (2011), on the other hand, uses their own word list adapted to financial contexts to build a measure of the tone of fundamental content in firms' 10-K filings. They find that firms with a high fraction of negative words is associated with more negative abnormal returns for a short period around the filing date. Jegadeesh and Wu (2013) use a different approach to quantify the positive and negative tone. Specifically, they rely on the market reactions to the 10-K filings to determine the strength and the direction (negative or positive) of each word, which later allows them to quantify the tone of the whole document. Similar to Tetlock et al. (2008) and Loughran and McDonald (2011), they find that such measure is informative of fundamental content as a more positive tone is associated with a higher return in the filing window. Other papers have produced similar results including Davis et al. (2006) and Demers and Vega (2008) among others.

Later studies that employ more sophisticated techniques such as Boudoukh et al. (2019), Heston and Sinha (2017) and Ke et al. (2019) also confirm that firm-specific news signals are informative and predictive of stock returns. Overall, most previous studies (except for Heston and Sinha (2017)) document return predictability from news signals for a few days after the news announcement. The most closely related to our study is Boudoukh et al. (2019). In this paper, the authors use a mixture of a rule-based information extraction platform and a support vector machine classifier to classify news that is relevant to firms' fundamental values, i.e, news that can be identified as belonging to one of the pre-defined categories of events (acquisition, analyst recommendation, deals, employment, legal, etc.) and those which is not. They show that days with identified news are associated with higher return volatility compared to those without news and those with unidentified news. Additionally, the fact that conditional on identified news, return volatility is similar during trading and overnight hours proves the importance of public information as a source of return volatility.

In the present paper, we use structured data processed by the AI platform of Causality Link to extract the informational content of news. Our analysis has two distinguishing features. First, instead of considering a unique source of information (as in Boudoukh et al. (2019)), we introduce a simple way to aggregate daily fundamental news across various sources available in the market (news stories, broker reports, call transcripts, etc.). The advantage of this aggregate signal is that it can measure not only the positiveness/negativity of the news content on a given KPI in a given document but also capture how widespread such content is in the market overall. Said differently, a largely reported news that appears in different news sources will be considered more important than if reported in only one single news source. Additionally, we also include in our analysis the universe of stocks available on the CRSP stock returns database and not just restrict it to only large-cap firms. This allows us to investigate heterogeneous market reactions for stocks of different market capitalization. We show that our signal is informative about firm-specific fundamental news and that a long-short strategy produces positive and significant abnormal returns till one day after the publication date. This finding is robust among different firm sizes. Second, another valuable characteristic of the Causality Link platform is that it is able to identify the tense of the news (whether the news release concerns something happening in the past, present or future) as well as the horizon of news (whether it relates to the near or the far future), a feature that has not been investigated before. This allows us to study the potential differences in the market reactions to news statements of different tenses and horizons. We show that news about the present or future receives higher market reactions compared to those about the past. Additionally, the market reaction to future news is mainly driven by near rather than distant future news.

Our study is also related to the growing literature on how the market reacts to ESG news. As there are only a few papers looking into the question, the evidence so far remains inconclusive. For example, Krüger (2015) shows that investors react strongly and negatively to negative events concerning a firm's corporate social responsibility (CSR). However, the study also documents some weak negative market reactions to positive CSR events. In his explanation, this may be due to positive CSR news concerning firms with a history of poor stakeholder relations.

Capelle-Blancard and Petit (2019), document a slight but significant drop in firms' market value when facing bad ESG news. On the contrary, firms facing good ESG news experience no significant change in their market value. Similarly, Cui and Docherty (2020) also document negative abnormal announcement returns for bad ESG news but not for good ESG news. Also, Serafeim and Yoon (2022) find that the stock market reaction to ESG news is moderated by the consensus ESG rating. In particular, the reaction is weaker when there is high disagreement among raters. ESG news have also a strong impact on analysts' forecasts. Derrien et al. (2021) show that after ESG incidents, analysts significantly downgrade their earnings forecasts and that ESG incidents affect earnings forecasts at longer horizons than other types of corporate incidents.

On the contrary, Aouadi and Marsat (2018) find that ESG controversies enhance value for firms that have high CSR scores (supposedly, large firms, located in countries with greater press freedom, followed by more analysts and have a good reputation overall). In our study, our news signal also enables us to categorize positive and negative news regarding ESG issues. Our results suggest that there are small but positive and significant market reactions to both positive and negative news around the publication date. Our analysis, by distinguishing the various types of news, is also able to compare the magnitude of market reactions to financial and ESG news. We show that ESG news trigger smaller reactions than financial news.

Our results have implications for practitioners. Financial analysts can use our approach to identify how news they create get incorporated into stock prices and what type of news they should pay most attention to. Moreover, investors may use our approach to set up trading strategies given that the market, for some stocks and some types of news, does not immediately reflect the informational content of news. In fine, the appeal of such trading strategies depends on whether their gross profitability is large enough to compensate for transaction costs. We leave this avenue of research for future investigation.

The rest of the paper is organized as follow. Section 2 discusses the data used and the construction of our aggregate news signal. Section 3 presents the empirical analysis on stock market reaction to news. Section 4 concludes.

2 Data, news signal and methodology

The AI platform of Causality Link has collected and processed tens of thousand articles everyday[1] from more than eighth thousand news sources and in twenty-seven languages. Unstructured data from broker research, earnings call transcripts, Edgar filings and in-house research is collected and transformed into structured data through Causality Link Natural Language Processing (NLP) pipeline. As each piece of content is added to Causality Link corpus in real time, it is queued for immediate parsing by the AI system, which uses both machine learning and symbolic techniques to recognize and understand different elementary concepts, particularly the Key Performace Indicators (KPIs) and Trends, described below. The processed content covers more than 4,000 companies[2] and over 1,600 industries and sectors. As the metadata is interpreted and collected from content, a streaming process indexes the thousands of individual mentions by the coordinated universal time (UTC) date and hour they were published. In this paper, we restrict the data to be in the period from January 1st, 2014 to December 31st, 2021.

Structured data from news can be identified based on KPIs. KPIs are components of the indicators that represent all numeric data mentioned in the news statements. Causality Link database contains 1,700 KPIs including financial indicators (e.g., profit, assets, liabilities, market share, etc.) and ESG indicators (e.g., carbon emission, human rights, etc.). The set of KPIs is a part of a large ontology that was specifically developed for this plat-

[1] Their system has processed more than 112 million texts as of January 21st, 2021.

[2] The number of companies includes both parent companies and their subsidiaries. The database covers news about 37,835 firms with unique trading tickers.

form. Each KPI is associated with a description including the associated country, industry, company and product. The ontology also provides coherence checks between the KPIs and the associated entities. For example, impossible indicators such as "the EBITDA of France" or "GDP of Apple" are avoided thanks to these checks.

On the time dimension, the system can also identify the tenses of the statements made in documents, i.e., whether the statement is about a change in KPI that occurs in the past, the present or is expected for the future. This information is stored in the attribute *offset* that can theoretically take any integer value. For example, an *offset* equal to -10, 0 or 10 means that the statement is about a change in a certain KPI that occurred ten days ago, at the moment or a projection/a plan for the next ten days, respectively.

In the following analysis, we focus on statements from news that capture the change in KPIs that are associated with specific company names. Each KPI is described by the attribute $favorability$ in $\{-1, 1\}$. $favorability$ takes a value of -1 if the KPI is associated with a negative impact on firms' stakeholders and takes a value of 1 if it is associated with a positive impact. For example, KPIs such as *taxes*, *operating expense*, and *carbon emissions* have a favorability that equals to -1 while KPIs such as *revenue*, *production*, and *sustainability* have a favorability of 1[3] The system is also able to identify the direction of change in KPIs mentioned in the statements and assigns a value in $\{-1, 0, 1\}$ to the variable *direction* for an increase/improvement, no change or a decrease in the KPIs, respectively. We define $trend = favorability \times direction$, a measure capturing the positiveness/negativity of a given statement for the firm. For example, *trend* will equal 1 if the statement is about an increase in this year's profit or about a decrease in operating expenses. Alternatively, if the statement is about a decrease in market share or an increase in carbon emissions, *trend* will equal -1. A statement that says the revenue in this quarter remains unchanged compared to last year will have a neutral trend, i.e.: *trend = 0.*

Below is an example of how the above contents are extracted from a news statement:

> "Despite the slowdown, both Google and Windows devices saw K-12 unit growth during the *fourth quarter, while* <u>*Apple's iPad volumes declined*</u> *year-over-year."* (USA TODAY, at 11:30:00 PM March 25, 2018)

From this sentence, the system identifies *Apple* as the company, associated with the KPI *transaction volume.* It also detects a *decrease* in Apple's transaction (or sales) volume ($direction = -1$). As this KPI is positively

[3] A few KPIs that have $favorability = 0$ (for example: *salary*, *quantity*, *common stock*, *inflation target* and *installed capacity*) are discarded from the sample as we are only interested in KPIs that have either negative or positive impacts on firms because this information is considered the most informative about firms' fundamentals.

related to the firm's shareholder value ($favorability = 1$), a decrease in its value is considered as bad news for Apple ($trend = -1$). Notably, the statement was made on March 25, 2018 but concerns the fourth quarter of the previous year. It is thus recognized as a movement in KPI that took place 83 days in the past ($offset = -83$). Similarly, for the same sentence, the system also identifies the company *Google* and the concerned KPI *unit growth*.

On a given day and for a given firm, the system can process different statements made in the same or in various documents. The trends of new statements can be the same (all negative or all positive) but can also be in different directions (some are negative, some are positive). To avoid looking-ahead bias, news on day t is the one that arrives between 4 pm (stock exchange closing time) on day t-1 and 4 pm on day t. We aggregate the trends at a firm-date level by defining the *positive trend percentage (ptp)* measure:

$$ptp_{i,d} = \frac{1 + n^{+}_{i,d}}{2 + (n^{+}_{i,d} + n^{-}_{i,d})} \tag{1}$$

where $n^{+}_{i,d}$ and $n^{-}_{i,d}$ are the number of statements that have positive and negative trends for firm i on day d, respectively. The *ptp* measure will take values in $(0, 1)$. If the number of positive and negative statements are equal or if all the statements are neutral, *ptp* will be equal to 0.5. Alternatively, *ptp* will be higher than 0.5 if there are more positive statements than negative ones and be lower than 0.5 in the opposite case. Accordingly, $ptp_{i,d}$ can capture how positive the news about a given firm is on a given day. The nature of our news signal is different from those seen in past studies. Instead of quantifying the tone of each article into a continuous variable (for example, Tetlock (2007), Ke et al. (2019), Loughran and McDonald (2011)), we first identify the trends of each news statement regardless of the magnitudes and then aggregate these news statements across all their appearance in the market on the same day. Therefore, a stock that has high *ptp* means that not only its news content is mainly positive but the positive mentions largely outnumber the negative ones on that day.

In the baseline analysis, *ptp* is constructed using *trend* of news statements for all types on news (i.e, all KPIs) and all tenses (past, present and future). We remove microcaps stocks, defined as those with market cap below \$300M, from our investment universe. Microcaps stocks (usually defined as those with market cap below \$300M, see, e.g., the SEC, or as those with market cap lower than the 20th percentile of market equity or NYSE stocks) can be influential in an equally weighted portfolio for two reasons as pointed out in Fama and French (2008). First, these stocks account for 60% of the total number of stocks on the market although only make up for about 3% of the total market cap on the NYSE-Amex-NASDAQ universe. Second, many of the anomalies in the literature are most prominent among microcaps. Hou et al. (2020) found that 64% to 85% of the anomalies in the literature in finance and accounting become insignificant when microcaps stocks are excluded from the sample. However, microcap stocks are difficult to trade for investors due to liquidity issues.

As news concerning different categories of KPIs may have different implications on stock returns, we also construct *ptp* using the sub-samples of data that correspond to KPI categories. Particularly, since news about environmental, social and governance (ESG) issues is quite specific and may receive different reactions from the market compared to financial news, we consider *ptp* signals constructed on the sub-samples based on ESG KPIs and financial KPIs separately.

As the information contained in news statements about the past or present can be partially incorporated into stock prices at the time the statements are made (for example, if the information was already available to some market participants, even if not officially published), the market reaction to such statements may be limited. Conversely, if the statements are about future events, this should be considered as "fresher" information to the market and one can expect reactions of greater magnitudes. Accordingly, we also construct *ptp* for all KPIs on the sub-samples based on the $offset$ value of each statement, where $offset < 0$, $offset = 0$ and $offset > 0$ indicates the statement is about the past, the present or the future, respectively.

Finally, there is evidence in the literature that market reactions to news may depend on firms' size. Information regarding smaller firms is more opaque due to less analyst coverage and therefore news should have a larger impact on prices. Our non-microcaps sample includes firms whose size range from very large to mid-cap. To test for a size effect, we also construct *ptp* signal on sub-samples based on market capitalization.

Data on daily stock returns are downloaded from CRSP and then merged with *ptp* data by dates and firms' tickers. Table 1 provides the descriptive statistics on the *ptp* measure calculated using the full sample and sub-samples constructed by aggregating news according to (1) the type of information released (financial news or ESG news), (2) the tense of the news (past, present or future) and (3) firms' size. Because we are interested in how information is incorporated into prices, only *ptp* signals constructed on trading days (weekdays except for holidays) are included. Our full sample covers 1,701,789 *ptp* signals that were built on 4,460 US stocks over the period January 2014 to December 2021.

Table 1 shows that financial news dominate the market. Across news types, tenses and firms' size, the means of *ptp* are slightly over 0.5 (except for ESG news) and all significantly different from 0.5 at 1% significant level, which suggests the news statements captured by Causality Link's NLP tool are more likely to be positive. Interestingly, ESG news signals are more likely than financial news to be positive (average *ptp* of 0.623, significantly different from 0.5 and from the average *ptp* for financial news at 1% significant level). We note that the majority of news signals concerns financial news statements and those concerning the present. The last row of Panel C suggests that stocks that belong to the Russell 1000 index are widely covered on news while mid-cap stocks, i.e., those that are not in the Russell 1000 index, receive much less attention in our news sources.

For each trading day, we sort all stocks according to their *ptp* signals. We then go long on the equally weighted portfolio that includes all stocks that have *ptp* equal or above the 90^{th} percentile and go short on the one that includes stocks with *ptp* equal or below the 10^{th} percentile. [4]

To observe whether the *ptp* signal is informative and how information from news is incorporated into prices, we look at portfolio returns within a period from minus five days to plus five days around day 0 - the day when the news is published on the market (also corresponding to the day the portfolio is constructed). Returns are then averaged across all daily portfolios throughout the period from January 1st, 2014 to December, 31st 2021. Table 2 gives the descriptive statistics of the portfolios constructed using the full sample and sub-samples based on news type, tenses and sizes. Overall, there is a clear discrepancy in the average value of *ptp* signals between the long and short portfolios (0.861 versus 0.285, significantly different from each others at 1% significant level). This gap is however smaller for the portfolios built on ESG signals (0.823 versus 0.311, significantly different from each other at 1% significant level). This suggests that our portfolios are constructed on strong news signals. The average number of stocks ranges from 15 to 123 stocks depending on the samples, with an average of 70 stocks, which leaves sufficient room for diversification. The portfolios built on our daily ranking of news signal involve a high daily turnover. In most of our sub-samples, the turnover is larger than 80%[5]. Notice that a 100% turnover means that one has to liquidate completely his portfolio formed on day $t-1$ to form a new one on day t. These high levels of turnover are similar to the ones documented in Ke et al. (2019).

Finally, we compute the returns of our long, short and long-short portfolios, as well as abnormal returns for each portfolio, computed as alpha relative to the Fama-French three-factor model. Daily returns of the three US factors (market, size and value) are retrieved from Kenneth French's data library.

[4] Due to the small number of stocks associated with ESG news, for the ESG subsample, besides the above way of portfolio construction, we also go long on the portfolio that includes stocks with *ptp* more than 0.5 and go short on the one that includes stocks with *ptp* less than 0.5.

[5] We assume that portfolios can only be constructed daily at the end of the day, after all signals on the market are observed. Portfolios are then re-balanced at the end of the following trading day.

3 Main results

3.1 Performance of portfolios based on all news

Table 3 and figure 1 present the performance of the long portfolio, the short portfolio and of the long-short strategy constructed using the aggregate *ptp* signal. Auto-correlation between daily portfolio returns can cause a downward bias in the conventional estimator of standard errors. To adjust for such potential bias, we apply a correction factor to the conventional standard errors as in Bence (1995). We describe in more detail the adjustment procedure in the appendix.

We observe an average daily return of -0.54 percentage points on day 0 for the short portfolio and an average daily return of 0.73 percentage points for the long portfolio built on *ptp* signals. These magnitudes are not only statistically but also economically significant. The long-short strategy offers a daily return of 1.27 percentage point. The magnitudes of average returns are comparable to the ones reported in Heston and Sinha (2017). Notice that average returns on the long, short and long-short strategies are also positive and significant (with smaller magnitudes) one day preceding the announcement day. These early market reactions are also observed in Heston and Sinha (2017) and Ke et al. (2019). Such a pattern can be explained by the fact that some investors have access to private information and trade on it. This pattern can also be due to the lag in publication from some news providers. For example, there may be some delay between the time when fundamental news initially arrives in the forms of companies' call transcripts or brokerage reports and the time when this information shows up in news stories (corresponding to the time at which we form our portfolios). We explore further this question in a later section of this paper.

News is incorporated very quickly into prices after it has been published. Day 1 average return (0.04 percentage point) on the long-short portfolio is smaller in magnitude compared to day -1 and day 0. Notice that in this analysis, we assume that the soonest time at which trades can take place is at the end of the publication date. Since some news may be published after the previous day close, we could obtain a higher profit if trading at the beginning of the day is allowed. For example, in the morning of date t, one can trade on news that is observed after 4:00 PM of date t-1. As in Heston and Sinha (2017), significant returns hardly reverse at least in the very short term, which confirms the informational content of our news signals. In an additional analysis, we also split the sample into sub-samples by year. Figure A1 in the appendix gives a snapshot of the performance of the long-short strategy from 2014 to 2021. The result confirms the robustness of the informational content of the *ptp* signals.

Table 4 reports the abnormal returns (alpha) of the long portfolio, the short portfolio and the long-short strategy relative to the Fama-French three-factor model. Several interesting facts emerge from the results. First, portfolio returns of the long-short strategy have relatively low exposures to the aggregate risk factors. The factor models can explain only 5.91% and 25.29% of the variation in the average returns of this strategy on day 0 and day 1, respectively. The positive returns on the long-short strategy on day 0 as observed in table 3 are mostly entirely alphas (1.27 percentage point).

Second, day -1 alpha for the strategy is in smaller magnitude but still significant. This finding is in line with the literature documenting positive, short-horizon auto-correlation for returns on portfolio of stocks (see Lo and MacKinlay (1988), Conrad and Kaul (1988) and Conrad and Kaul (1989), for example). Later study shows that partial price adjustment (traded prices of some stocks do not fully reflect information held by traders) is the main source of auto-correlation in short-horizon stock return Anderson et al.. Additionally, the magnitude of alpha on the short portfolio is larger compared to the one on the long portfolio. This can be explained by the fact that investors often face short-sale constraints, which prevent them from fully exploiting bad news. Therefore, comparing to good news, bad news is absorbed into prices more slowly.

Third, average returns on the short portfolio are more heavily loaded on the HML and SMB factors compared to those on the long portfolio. A possible explanation is that the short portfolio may contain more value and small stocks, more likely to be in distress and thus experiencing large negative returns when negative news occurs.

Overall, the performance of portfolios built on the aggregate *ptp* signals shows strong market reactions to news around day 0. This suggests that the *ptp* signals are successfully capturing the aggregate tone of news on the market. Additionally, it also appears that though new information is absorbed into prices fairly quickly, there is still possibility to make profits from a simple trading strategy based on these signals.

3.2 Tenses and horizons of news

Fundamental news regarding the past, present or future may be incorporated into prices in different manners. For example, information regarding future events may contain more "surprises" for the market compared to the one regarding past events as the latter may be well expected before the official publications. Therefore, the magnitudes of the market reactions to news signals around day 0 should be larger for news about the future relative to those regarding the past or the present.

Table 5 and figure 2 show the average returns on the long, short and long-short portfolios. Overall, the majority of the market reactions to the three types of news take place on day -1 and day 0. As expected, the average returns are of higher magnitude for news regarding the future relative to those regarding the past or the present. Specifically, the average returns on the long-short strategy on day 0 is 1.26 percentage points higher for the news about the future compared to those about the past, and 0.54 percentage points higher compared to those regarding the present. This discrepancy comes from both the difference in reactions to good news and the ones to bad news.

Table 6 shows the abnormal returns of the portfolios built on news signals of different tenses. As one may observe, alphas of daily portfolio returns on day 0 are positive and significant in all of the three sub-samples. We also observe higher magnitudes for alphas on day 0 in the future sub-sample compared to the other two ones. Alphas of day 1, on the other hand, are only significant for present and future news, which indicates that they are more difficult for the market to digest.

On a given day and for a given firm, there may be news statements regarding different tenses. The observed market reactions might reflect the combination of all news statements instead of those of a single tense. In an attempt to separate the impact of statements of different tenses on stock returns, we also add a constraint to the long and short portfolios: on the same day, a stock i cannot appear in more than one portfolio. In other words, for each tense-based portfolio, we eliminate stocks that are the most likely to be affected by news concerning other tenses. Table A1 and figure A2 in the Appendix show that there are still significant market reactions to news on day -1 and day 0 although they are of smaller magnitudes. The long-short strategy based on future news still show the highest returns on day 0 compared to the ones based on past and present news. Day 1 abnormal returns on the long-short strategy reported in table A2 are of smaller magnitudes and no longer significant (except for the long-short portfolio built on news of the present tense).

News regarding near future cash-flows should have a greater impact on prices compared to those regarding distant future cash-flows for at least two reasons. First, the change in investors' expectation caused by news is discounted less heavily in the former case. Second, news regarding cash-flows in the distant future is likely to be more uncertain compared to those regarding cash-flows in the near future. We therefore divide the sample of future-tense *ptp* signals into near and distant future sub-samples. As mentioned in Section 2, the attribute $offset$ in our data set provides the exact time in the future that the news statements are about. The median of the offset for future news is 60 days. We use this value as a threshold to distinguish between near and distant future news.

Table 7 and figure 3 confirm our prediction. Near future news triggers a much larger market reaction compared to distant future news. Specifically, the average return on the long-short strategy based on *ptp* signals that have offset values less than 60 days is 0.72% and 2.37% on day -1 and day 0, respectively versus 0.44% and 0.86% of those based on news having $offset$ larger than 60 days. Table 8 shows that day 0 abnormal returns of portfolios built on near future news are also higher compared to those of portfolio built on distant future news. Additionally, the long-short strategy on both near and distant future news can create significantly positive alphas on day 1 (0.05% and significant).

The difference in the market reaction to near and distant future news may come from the possibility that different horizons are associated with different types of news. We report in table A3 the list of top twenty KPIs appearing in near future news (offset ≤ 60) and distant future news (offset > 60) and their associated chance of appearance in all news statements in the sample. These top twenty KPIs account for more than 90% of all KPIs that can be identified in near and distant news, respectively. It is noticeable that the two lists of KPIs are very similar: they share 17 common items out of 20. Therefore, it is unlikely that different types of news are released regarding the near and distant futures.

The average *ptp* of the long portfolio of near future news is 0.792, significantly lower than that of distant future news (0.815). However, as shown in Table 7, the magnitude of returns on the long portfolio of near future news on day 0 is almost three times larger. This suggests that the horizons of future news should have an impact on the magnitude of the market reaction beyond the difference in average *ptp* or the type of news that are released. Firms may have a tendency to issue more positive news regarding the distant future because their credibility is more difficult to assess; anticipating this, market participants may discount them more more strongly.

3.3 News coverage

Big news such as a surprise in earnings or a large-scale layoff is expected to receive a strong market reaction compared to less significant news. One way to identify big news is to look at how many times news about a certain KPI is mentioned in the media. Also notice that news about big companies is often widely covered compared to those of smaller ones. In this case, the number of times that a piece of news is mentioned does not necessarily reflect its significance to the firms. Therefore, we consider in the following analysis only large-cap firms that belong to the Russell 1000 index as a way to control for firm size. Among these firms, big news should be relatively more widely covered compared to insignificant news. For each stock and each KPI, we count the number of new statements mentioning that same KPI on day 0. The median number of news statements per KPI is 1, which suggests that most of the news has low coverage. We divide the sample into four sub-samples according to the number of mentions for each KPI on day 0. Table 9 and figure 4 shows the market reaction to news with different coverage. Higher-coverage news in general triggers higher market reaction on days -1 and 0. Specifically, the average returns on the long-short strategy increase monotonically from 0.22% on day 0 for news with only one mention to 2.03% for those with more than ten mentions. Table 10 shows that the day 0 alphas also increase monotonically with the number of mentions. Particularly, alphas of the long-short strategy on very high-coverage news (those with more than ten mentions) are almost ten times larger compared to low-coverage news (those with only one mention).

3.4 News novelty

We expect that novelty affects the way the market incorporates news: fresh news are expected to move markets more than stale news on the day the news is released. Moreover, we have argued that significant portfolio returns on day -1 might be due to either the delays in publications from some news providers or the fact that some investors have access to private information. Looking into sub-samples of news based on its novelty would allow us to shed light on the most likely explanations. In particular, if the early market reaction is due to investors' private information, we could continue to observe early market reactions in the sub-sample of fresh news. If the effect is due to news' staleness, the early market reaction should disappear when focusing on fresh news.

To account for the novelty in daily news, we construct the variable *novelty* for the KPI k and firm i as follow:

$$novelty_{k,i,d} = \frac{S_{k,i,d}}{\sum_{j=d-99}^{d} S_{k,i,j}}$$

where $S_{k,i,d}$ and $S_{k,i,j}$ are the total number of news statements involving the KPI k of firm i on days d and j, respectively. Our measure of news novelty will be close to 1 if in the past one-hundred days (day $d-99$ to day d), the news statements about the KPI k of firm i mainly occur in day d. Alternatively, if the majority of news occurs from day $d-99$ to day $d-1$, *novelty* will be close to 0. Accordingly, the closer *novelty* to 1, the "fresher" the news is.

In the following analysis, we define "fresh" news as those having *novelty* greater than 0.5, i.e, the majority of the news statements regarding firm i and KPI k appear on day 0. On the contrary, "stale" news is defined as those having *novelty* lower than 0.5, i.e, the majority of the news statements regarding firm i and KPI k appears before day 0. Table 11 and figure 5 contrast the market reactions to fresh and stale news.

A expected, we observe that the magnitude of portfolio returns is much higher for fresh news compared to stale news, on day 0 at which the news is published. The daily return of the long-short portfolio based on fresh news is 1.77% on day 0, while it is only 1.20% for stale news. On the contrary, long-short portfolio return on day -1 is 0.32% (significant) for fresh news which is less than half the return of 0.81% (significant) for stale news. This results indicates that the day -1 price reactions we observed so far may be due both to private information, because the returns on day -1 for fresh news are significant, and to delays in publications by some news providers, because the return on day -1 for stale news is much larger than for fresh news. Table 12 confirms the difference in the magnitude of alphas between the two types of news. Without surprise, portfolios built on fresh news give much larger alphas compared to those built on stale news, both at date 0 and at date 1. This finding is consistent with Ke et al. (2019) in which they also report much higher portfolio returns after the announcement of fresh news compared to stale news.

3.5 Signal based on fundamental news only

Our construction of *ptp* on all KPIs may include both fundamental news as well as news regarding past price changes of a particular stock. In the former case, fundamental news usually reflects changes in firms' expected cash-flows and therefore induce investors and analysts to update their stock valuation. In the latter, it is also common for news stories to cover changes in stock prices (which are often mentioned in the past or present tense). These news stories might also cause market reaction following the news, for example if investors form beliefs about future stock market returns by extrapolating past returns (De Long et al. (1990), Hong and Stein (1999) and Barberis et al. (2015)).

In the following analysis, we focus on purely fundamental news. We compute *ptp* based on news of all KPIs except for those concerning the KPI "stock price" mentioned in the past and present tense. We are left with 1,610,148 *ptp* signals in this sub-sample. Figure 6 and table 13 show the market reaction to this type of news. The raw day 0 average return on the long-short strategy is 0.45% and significantly different from 0. However, the magnitudes of portfolio returns on day 0 are much lower for fundamental news compared to those of all news (0.45% vs 1.27%). Table 14 presents the abnormal returns on day 0 and day 1 for fundamental news. As expected, we also observe a small (but significant) alpha for day 0 on the long-short strategy. The alpha on day 1 returns is only marginally significant.

3.6 ESG and financial news

Table 15 and figure 7 show the performance of portfolios constructed using the *ptp* signals on financial news and ESG news.

Financial news receive stronger reactions compared to ESG ones. The magnitudes of returns on portfolios built on the *ptp* signals of financial news are much larger than those of ESG news (0.75% versus 0.09% for day 0 returns on the long portfolio and -0.55% versus 0.09% for day 0 returns on the short portfolio).

On the performance of portfolios built on ESG KPIs, three important observations stand out. A first observation is that there is very weak market reactions to ESG-related good news compared to news related to other KPIs. The average returns of the long portfolios are 0.08%, 0.09% and 0.06% on day -1, 0 and 1, respectively (all significantly different from zero). This observation is robust to different portfolio construction methods. In Panel C of table 15 and 16, we form the long and short portfolio by buying the stocks whose *ptp* is equal or above the 0.5 and sell those whose *ptp* is equal or below the 0.5 percentile. Since the number of stocks that have signals regarding ESG is low, this way of forming portfolios would allow more room for diversification. The performance of this strategy is very similar to those reported in Panel B of table 15 and 16.

One reason for the low-magnitude returns could be either that some investors simply discard ESG related news. Often qualitative by nature, this information is more tricky to be incorporated into firms' valuation. Moreover, there is no news announcement schedule, and no analysts or macroeconomic forecasters offering their predictions to market participants. Another explanation could be that ESG news may concern cash flows in the very distant future and discounting would induce a small impact on stock prices. We check and confirm this is the case in our data set. For future-tense news, the one regarding financial KPIs has an average offset value of 196 days, i.e, on average, future news discusses the movement in financial KPIs that happens in the next 196 days while future news regarding ESG KPIs has an average offset value of 1,325 days (the difference is significant, p-value = 0.00).

Financial news often contains both news regarding the past movement of stock price and fundamental financial news, while ESG news is purely fundamental. Therefore, the distinction in the magnitudes of the market reactions to the two types of news may come from the difference between the market reactions to stock-price related news and the ones to fundamental news. In figure A3 (appendix), we also contrast the magnitudes of ESG news with fundamental financial news (where stock-price news is not included). It is clear that the average returns on the long-short strategy for ESG news are of much smaller magnitudes compared to those for fundamental financial news. This confirms that the nature of the two news types also drives the difference in market reactions.

A second important observation is that we observe positive (although of very small magnitude) market reactions to ESG-related bad news. The average returns on the short portfolio on days -1, 0 and 1 are 0.08%, 0.09% and 0.06%, respectively, all significantly different from zero. This result is also robust for the other way of portfolio construction (Panel C). One explanation for our results is that investors simply react positively to any ESG news including those which are positive, neutral or negative. Positive reaction to all types of news may come from the fact that investors' attention is driven towards these stocks, leading to more frequent purchases (see Barber and Odean (2008)). Figure A4 in the appendix looks at the performance of the three portfolios based on *ptp* signals. The "bottom" portfolio consists of stocks that have *ptp* less than or equal to the 25^{th} percentile of all *ptp* in the same day. The "top" portfolio consists of stocks that have *ptp* signals more than or equal to the 75^{th} percentile. Finally, the "middle" consists of stocks that have *ptp* within the inter-quartile. We observe positive and significant average returns on day -1, 0 and 1 in all three portfolios. This confirms our previous conjecture that the market seems to react positively to both positive and negative ESG news. A consequence of the similar performance of the long and short portfolios is that the long-short strategy based on ESG news does not produce significant returns, even on day 0.

Table 16 presents the abnormal returns for the portfolios constructed on sub-samples based on financial and ESG KPIs. The alphas of the long-short strategy built on financial KPIs are of much higher magnitudes compared to those of portfolios built on ESG ones: 1.29 percentage point (significantly different from 0) versus -0.01 percentage point (insignificantly different from 0) for financial KPIs and ESG news on day 0 respectively. The long-short strategy on financial news also generates positive and significant alpha (0.04%) on day 1.

Previous studies exhibited contrasted results on the impact of ESG news. For example, Capelle-Blancard and Petit (2019) show that negative ESG events trigger a drop in firms' market value while positive ESG events do not have an impact. Krüger (2015), on the other hand, finds that both positive and negative events related to CSR issues create a drop in firms' market value.

Figure A5 in the appendix investigates further the evolution of the market reaction to ESG news from 2014 to 2021. The result suggests that ESG news impact is relatively stable over the sample period. In particular, there was no increase in the impact of ESG news through time.

A third important observation is related to the interaction between ESG and financial news. As the majority of news on the market is financial news, it is likely that ESG news is released on the same day that financial news is released either by coincidence or by firms' strategy. If this is the case, the market reactions to ESG news observed earlier may be affected by financial news that comes in the same day. Table A4 and A5 show the market reactions to ESG news where there is no extreme financial news in the same day. Specifically, we only keep stocks that appear in the long or short portfolios built on ESG signals and do not appear in the portfolios built on financial news on the same day. In other words, we eliminate from the ESG portfolios the stocks that are the most likely to be affected by financial news. We observe even weaker market reaction to bad ESG news. This suggests the positive market reaction to bad news observed earlier is probably a result of some good financial news that is released in the same day.

3.7 Different firms' size

News regarding large firms are more likely to appear in the media. Additionally, firms with large market capitalization also tend to have higher analyst coverage, which makes it easier for investors to get access to fundamental information. As a result, small and insignificant news about these firms tend to show up more often and one can expect little market reaction to such news. To test this prediction, we split our sample into two sub-samples, one that includes Russell 1000 index stocks[6](referred to as *Russell 1000* sub-sample) and one that does not include these stocks (referred to as *Russell 1000-excluded* sub-sample). Since the Russell 1000 includes US firms with the largest market capitalization in the market, the sample of these stocks should be representative of large-caps firms.

Table 17 and figure 8 suggest that market reactions to news regarding stocks included in the Russell 1000 index are of much smaller magnitude compared to those for the rest of the market (0.67% versus 1.79% for day 0 average return on the long-short strategy for the Russell 1000 and the Russell 1000-excluded sub-sample, respectively). Table 18 shows that alphas for day 1 returns of the long-short strategy are only positive and significant for Russell 1000 excluded stocks and mainly driven by returns on the short portfolio. This observation can be explained by the fact that it is much harder for investors to short-sell smaller and less popular stocks compared to large and well-known ones. This consequently leads to a slow incorporation of bad news into prices for these stocks.

4 Conclusion

In this paper, we use news data extracted by the AI platform of Causality Link to build daily aggregate signals that aim to capture the tone of news, its tense and how prominent that news is on the market. We then test the informational content of this signal by looking at the market reaction to news around the announcement day. This allows us to examine when and how fundamental information from news is incorporated into prices.

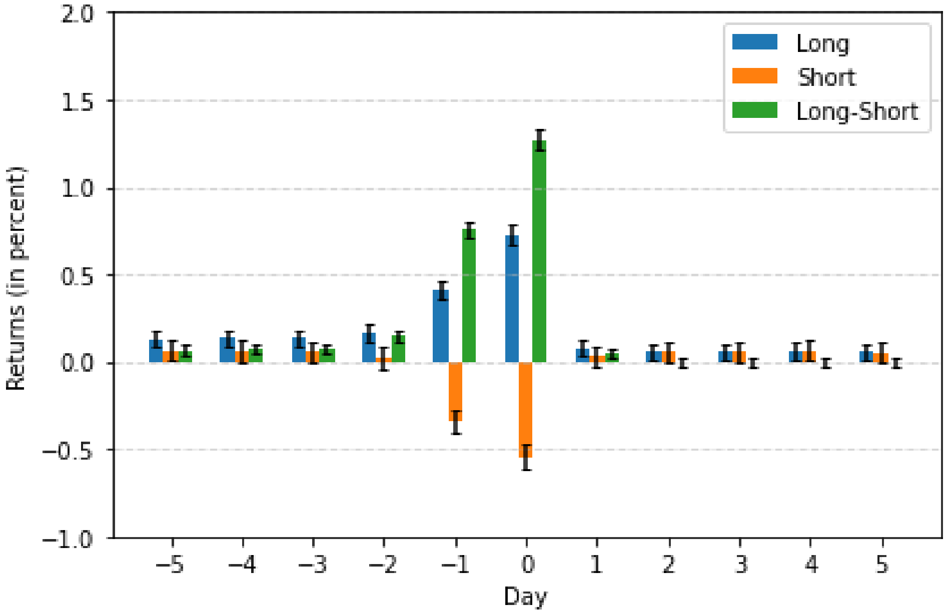
Particularly, each day, we build a long portfolio that consists of stocks that are above the 90^{th} percentile of the market on the same day, in terms of percentage of good news, and a short portfolio that consists of stocks that are below the 10^{th} percentile. We also form portfolios on the signals constructed using sub-samples by Key Performance Indicators (KPIs), tenses and firm sizes. By looking at the performance of such portfolios and of the long-short strategy, we offer the following results.

[6]As February 20, 2022.

First, the aggregate news signal can capture the content of fundamental news on the market. Second, there are some early reactions to news. Third, we find evidence of return predictability by the news signal and show that one can build a profitable strategy (abstracting from transaction costs). Fourth, there are stronger market reactions to news regarding the future compared to those regarding the present or the past. Fifth, high-coverage news trigger more market reaction compare to low-coverage one. Sixth, fresh news receives larger market reaction following its release but less early market reaction compared to stale news. Seventh, there are weak positive reactions to both positive and negative ESG news. Eighth, the strength of market reactions is larger for mid-caps than for large cap stocks.

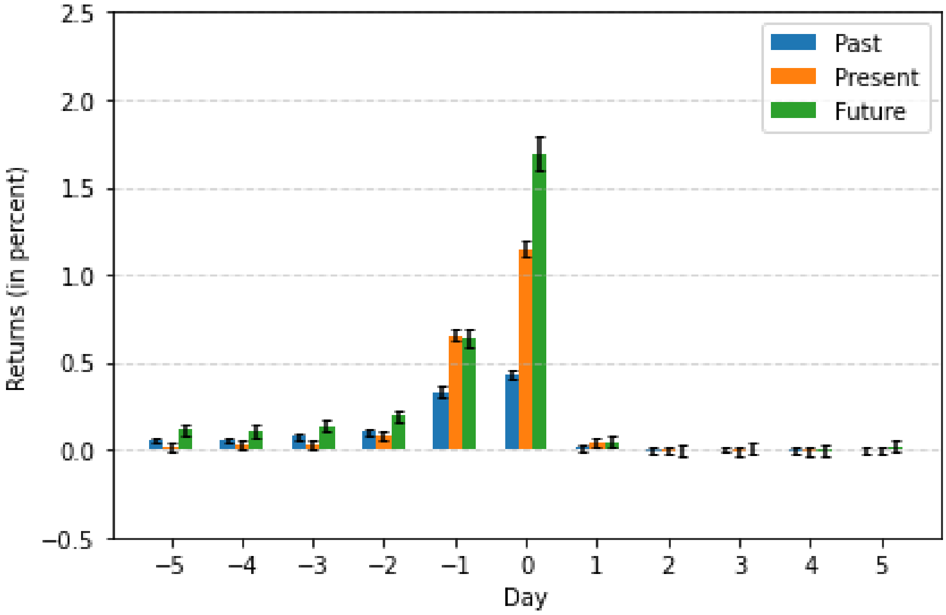
In future research, it could be interesting to study if our news-based investment strategy can generate profits after taking into account transaction costs. Another venue for future research is to study the impact of news uncertainty on investors' response to news. Further research is also needed to shed light on the observed positive market reactions to both positive and negative ESG news. If the majority of ESG news are from companies' own reports and transcripts, a positive reaction can be explained by the fact that investors think companies care about ESG issues even when the news released is bad. Therefore, it would be interesting to run a refined analysis, distinguishing the various sources of ESG news that appear in the data. Finally, since fundamental news from call transcripts or analyst reports take some time to appear in news stories, future research could study whether the long-short strategy based on news extracted from call transcripts are more predictive of future returns.

Figure 1: Return on the long portfolio, the short portfolio and the long-short strategy



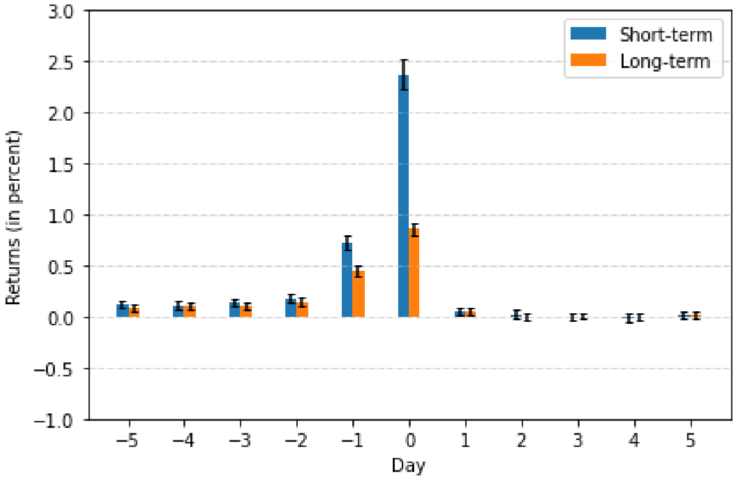
The bar charts present the average returns of the long portfolio, the short portfolio and the long-short strategy in the period [-5, +5] days around the portfolio construction days, on the full sample. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 2: Return on the long-short strategy for different tenses



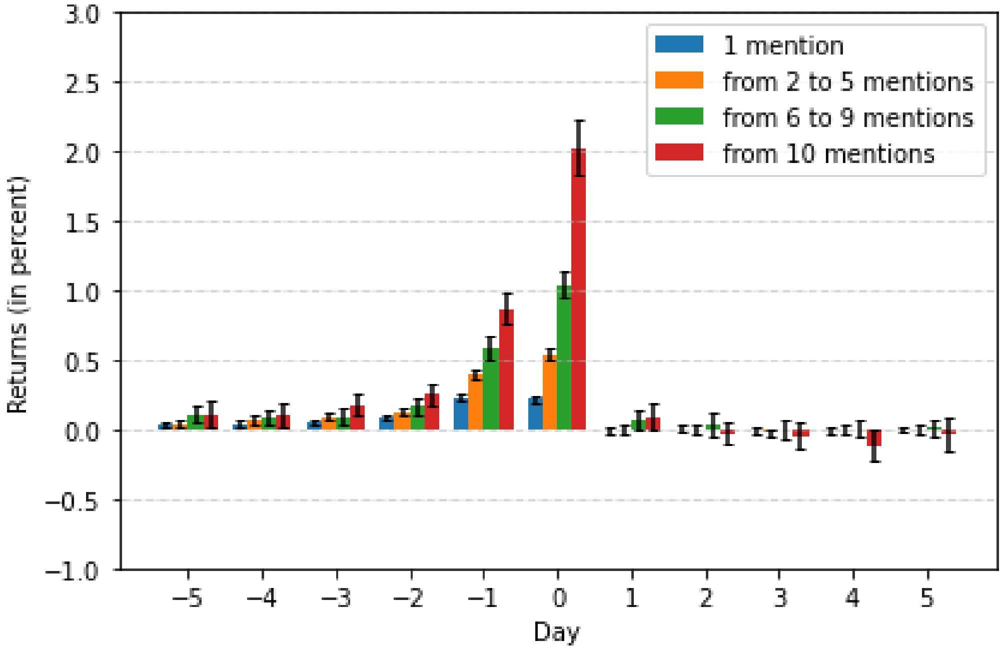
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning news about the past, the present and the future. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 3: Return on the long-short strategy for near and distant future news



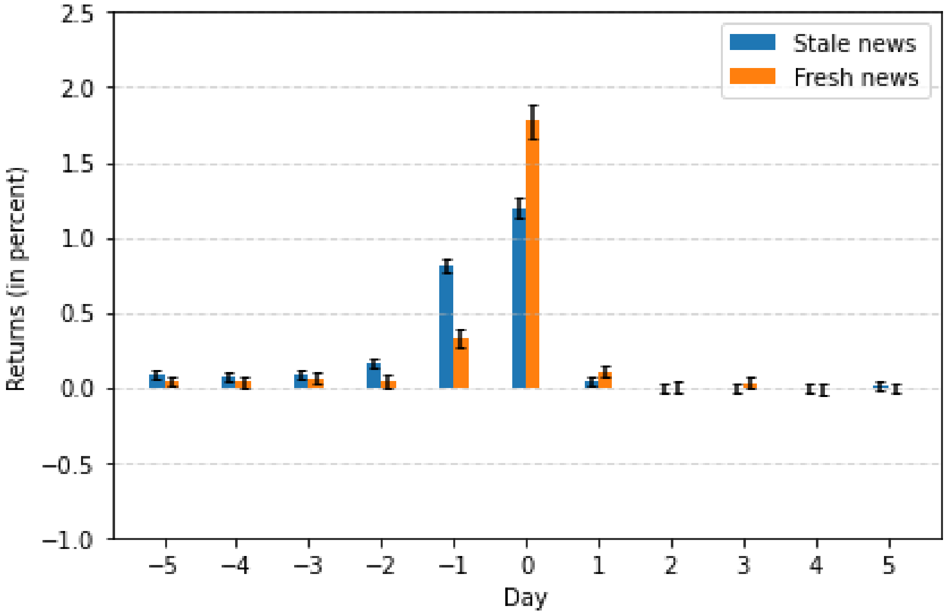
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples of near future news ($offset \leq 60$ where 60 is the median of positive offset) and distant future news ($offset > 60$). The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 4: Return on the portfolios for news of different coverage



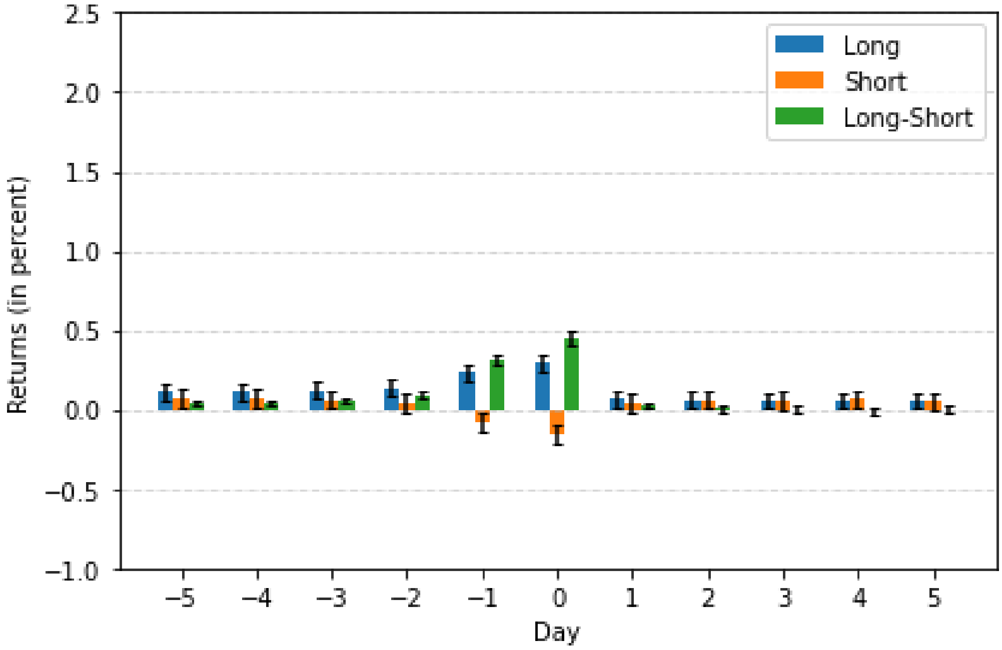
The bar charts present the average returns of the Long - Short strategy (for stocks in Russell 1000 index) for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning news coverage. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 5: Return on the portfolios for fresh and stale news



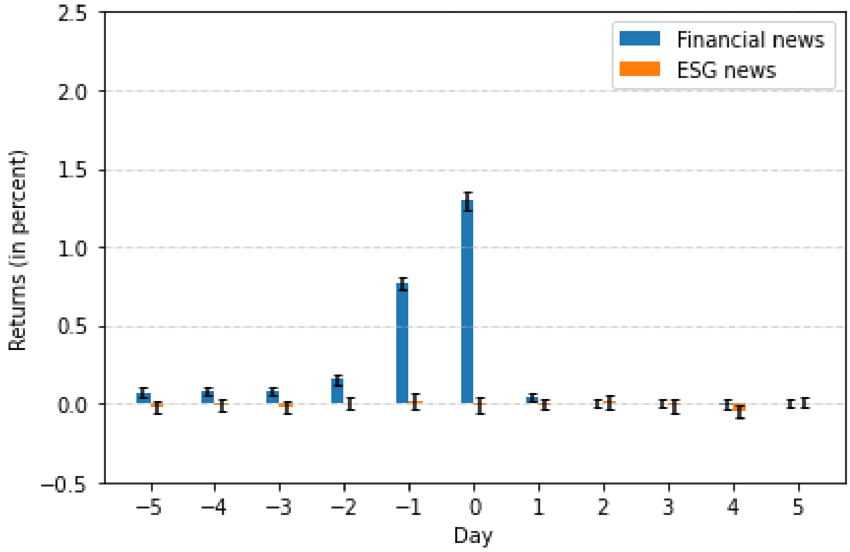
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning fresh and stale news. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 6: Return on the portfolios for fundamental news



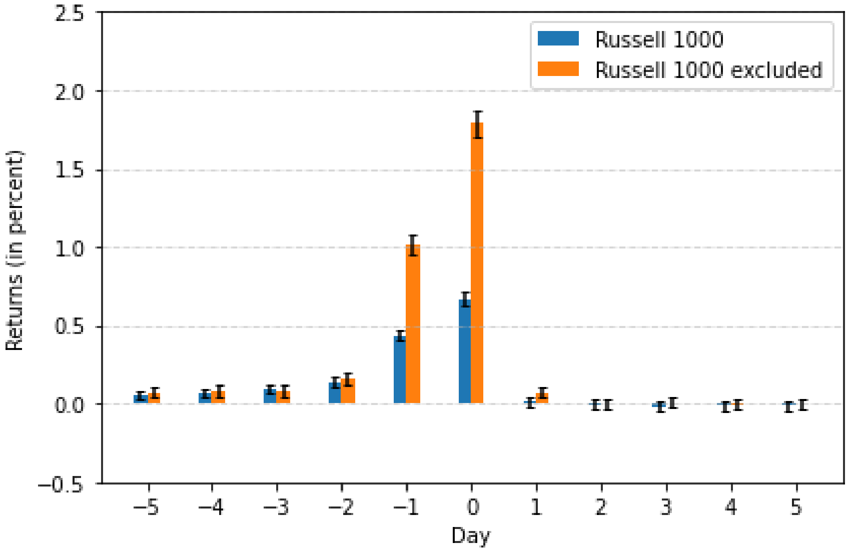
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning fundamental news. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 7: Return on the long-short strategy on ESG and Financial news



The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning ESG and financial KPIs. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure 8: Return on the long-short strategy for different firm size



The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples of Russell 1000 stocks and Russell 1000 stocks excluded. The y-axis is in percentage points. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 1: Descriptive Statistics of Daily News Signals

Panel A: Sub-samples of news based on the type of information released

KPI	All	ESG	Financial
ptp mean	0.596	0.623	0.592
ptp stdev	0.184	0.168	0.184
Nb ptp signals	1,701,789	212,959	1,677,235
Avg nb of signal per stock	381.57	62.52	376.06
Avg nb of signal per day	844.56	105.69	832.37
Nb of news	144,922,406	4,474,192	140,448,214
Avg nb of news per stock	32,493.81	2,220.44	31,490.63
Avg nb of news per day	71,921.79	1,313.62	69,701.35
Nb distinct stocks	4,460	3,406	4,460

Panel B: Sub-samples of news based on tenses

Tense	Past ($offset < 0$)	Present ($offset = 0$)	Future ($offset > 0$)
ptp mean	0.575	0.585	0.578
ptp stdev	0.181	0.176	0.182
Nb ptp signals	698,766	1,414,741	850,500
Avg nb of signal per stock	159.32	318.92	194.05
Avg nb of signal per day	346.78	702.10	422.08
Nb of news	35,522,499	90,359,006	19,040,901
Avg nb of news per stock	8,099.06	20,369.48	4,344.26
Avg nb of news per day	17,629.03	44,843.18	9,449.58
Nb distinct stocks	4,386	4,436	4,383

Panel C: Sub-samples of news based on firms' size

Company size	Russell 1000	Russell 1000 excluded
ptp mean	0.610	0.583
ptp stdev	0.183	0.184
Nb ptp signals	809,768	963,042
Avg nb of signal per stock	813.02	267.81
Avg nb of signal per day	401.87	477.94
Nb of news	71,629,280	73,293,126
Avg nb of news per stock	71,916.95	20,381.85
Avg nb of news per day	35,548.03	36,373.76
Nb distinct stocks	996	3,596

NOTE: In Panel A, *ptp* is computed using news statements in the sub-samples based on KPIs. In Panel B, *ptp* is computed using news statements in sub-samples based on tenses. In Panel C, *ptp* is computed using news statements in sub-samples firm sizes.

Table 2: Descriptive statistics of the equally weighted portfolios constructed on daily news signals

Panel A: Portfolios built on *ptp* in sub-samples based on the information released

	Long Portfolio				Short Portfolio			
KPI	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover
All	0.861	0.043	97.90	0.83	0.284	0.065	117.36	0.86
ESG	0.823	0.052	15.10	0.88	0.311	0.062	16.44	0.90
Financial	0.858	0.043	97.40	0.83	0.248	0.065	123.54	0.86

Panel B: Portfolios built on *ptp* in sub-samples based on tenses

	Long Portfolio				Short Portfolio			
Tenses	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover
Past	0.826	0.054	45.34	0.88	0.292	0.061	62.60	0.876
Present	0.844	0.044	85.65	0.82	0.293	0.060	109.11	0.86
Future	0.823	0.052	56.86	0.87	0.289	0.062	72.51	0.88

Panel C: Portfolios built on *ptp* in sub-samples based on firm sizes and stock prices

	Long Portfolio				Short Portfolio			
Size	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover	Mean ptp	Std Dev ptp	Avg no. of stocks	Turnover
Russell 1000	0.871	0.040	45.77	0.82	0.288	0.067	53.04	0.854
Russell 1000 excluded	0.859	0.042	48.35	0.86	0.285	0.065	76.67	0.865

NOTE: The long portfolios are constructed daily, by selecting stocks that have *ptp* signals above or equal to the 90^{th} percentile of the total signals in the same day. The short portfolios are constructed daily, by selecting stocks that have *ptp* signals below or equal to the 10^{th} percentile of the total signals in the same day. We also report the daily's turnover, defined as $\frac{1}{2T}\sum_{t=1}^{T}(\sum_i |w_{i,t} - w_{i,t-1}(1 + r_{i,t})|)$, where $w_{i,t}$ and $w_{i,t-1}$ are the weights of stock i in the portfolio on day t and $t-1$, respectively and T is the number of daily portfolios constructed.

Table 3: Returns on portfolios based on news signals arriving at date 0

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.13	5.37	0.06	2.05	0.07	4.71
-4	0.13	5.40	0.06	1.98	0.07	5.14
-3	0.13	5.44	0.06	1.87	0.08	5.66
-2	0.17	6.71	0.02	0.72	0.15	9.92
-1	0.41	14.89	-0.34	-10.21	0.75	33.97
0	0.73	23.44	-0.54	-15.76	1.27	42.74
1	0.08	3.20	0.03	1.10	0.04	3.20
2	0.06	2.37	0.06	1.89	-0.00	-0.02
3	0.06	2.28	0.05	1.80	0.00	0.06
4	0.06	2.47	0.06	2.08	-0.01	-0.36
5	0.05	2.25	0.05	1.73	0.00	0.04

NOTE: On every trading day, we sorted all stocks on the basis of *ptp* constructed on the full sample. We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolios in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 4: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model

Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.27***	-0.09***	-0.02***	-0.19***	5.91%
	(33.16)	(-3.21)	(-4.85)	(-3.88)	
L	0.67***	1.02***	0.10***	0.48***	76.83%
	(30.87)	(56.32)	(3.76)	(16.52)	
S	-0.60***	1.11***	0.30***	0.67***	79.62%
	(-22.91)	(56.26)	(8.94)	(17.59)	
			Day 1		
L-S	0.04***	-0.09***	-0.25***	-0.25***	25.29%
	(3.45)	(-5.92)	(-6.09)	(-6.98)	
L	0.01**	0.99***	0.09***	0.37***	31.55%
	(2.04)	(124.78)	(5.85)	(25.26)	
S	-0.03***	1.08***	0.35***	0.62***	90.48%
	(-2.70)	(66.68)	(9.95)	(21.31)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 5: Returns of portfolios based on news signals arriving at date 0 with different tenses

Panel A: News about past events						
	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.11	3.91	0.05	1.69	0.06	7.17
-4	0.12	4.33	0.07	2.19	0.06	6.83
-3	0.13	4.62	0.05	1.53	0.08	8.64
-2	0.15	5.14	0.05	1.45	0.10	11.47
-1	0.27	9.07	-0.06	-1.92	0.33	20.02
0	0.33	11.15	-0.10	-3.37	0.43	28.53
1	0.07	2.60	0.06	1.95	0.01	1.19
2	0.06	2.31	0.07	2.18	-0.00	-0.50
3	0.06	2.26	0.06	1.88	0.00	0.34
4	0.06	2.24	0.06	2.05	-0.01	-0.58
5	0.06	2.31	0.06	2.06	-0.00	-0.32
Panel B: News about present events						
	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.11	4.45	0.09	3.00	0.02	1.53
-4	0.12	4.86	0.09	2.94	0.03	2.49
-3	0.12	4.92	0.09	3.11	0.03	2.40
-2	0.14	5.69	0.06	2.06	0.08	5.67
-1	0.36	13.19	-0.30	-9.61	0.66	37.13
0	0.61	20.36	-0.54	-17.27	1.15	50.22
1	0.08	3.27	0.04	1.33	0.04	3.36
2	0.06	2.39	0.06	2.10	-0.00	-0.33
3	0.06	2.39	0.06	2.25	-0.01	-0.53
4	0.05	2.27	0.06	2.17	-0.01	-0.72
5	0.05	2.06	0.05	1.75	-0.00	-0.09

(Continued on next page)

Table 5 – *(Continued from previous page)*

Panel C: News about future events						
	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.14	5.57	0.03	0.85	0.11	7.01
-4	0.14	5.53	0.03	0.92	0.11	6.24
-3	0.14	5.74	0.01	0.33	0.13	8.19
-2	0.17	6.59	-0.02	-0.63	0.19	10.52
-1	0.41	13.61	-0.23	-0.68	0.64	24.54
0	1.03	26.57	-0.66	-16.72	1.69	34.81
1	0.08	3.26	0.04	1.19	0.04	2.73
2	0.06	2.41	0.06	1.92	-0.00	-0.04
3	0.06	2.32	0.05	1.60	0.01	0.49
4	0.05	2.22	0.06	2.02	-0.01	-0.52
5	0.06	2.40	0.04	1.25	0.02	1.17

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning news about the past (Panel A), the present (Panel B), and the future (Panel C). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 6: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model, of portfolios constructed based on news signals with different tenses

Panel A: News about past events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.43***	-0.05***	-0.17***	-0.08***	11.11%
	(24.32)	(-3.78)	(-7.13)	(-2.63)	
L	0.28***	1.00***	0.18***	0.46***	90.97%
	(21.22)	(97.23)	(9.70)	(26.35)	
S	-0.16***	1.06***	0.35***	0.54***	87.39%
	(-10.36)	(71.02)	(12.03)	(16.75)	
			Day 1		
L-S	0.00	-0.04***	-0.16***	-0.09***	20.49%
	(0.44)	(-3.50)	(-7.37)	(-5.57)	
L	0.01	1.00	0.19***	0.43***	95.48%
	(1.10)	(99.42)	(15.09)	(19.96)	
S	0.00	1.04***	0.35***	0.52***	91.51%
	(0.09)	(79.82)	(12.54)	(18.59)	

(Continued on next page)

Table 6 – *(Continued from previous page)*

Panel B: News about present events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.15***	-0.07***	-0.14***	-0.13***	3.91%
	(40.24)	(-2.79)	(-4.09)	(-3.42)	
L	0.55***	1.02***	0.13***	0.462***	78.25%
	(27.62)	(61.01)	(5.63)	(16.43)	
S	-0.60***	1.09***	0.27***	0.59***	83.34%
	(-29.91)	(60.31)	(9.52)	(20.10)	
			Day 1		
L-S	0.04***	-0.06***	-0.17***	-0.22***	15.68%
	(3.41)	(-4.38)	(-5.40)	(-7.47)	
L	0.02**	0.99***	0.12***	0.33***	89.96%
	(2.41)	(121.51)	(8.52)	(24.45)	
S	-0.02**	1.04***	0.28***	0.56***	90.85%
	(-2.44)	(68.56)	(11.33)	(18.16)	

(Continued on next page)

Table 6 – *(Continued from previous page)*

Panel C: News about future events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.69***	-0.10***	-0.23***	-0.05	2.24%
	(26.17)	(-2.71)	(-4.25)	(-0.82)	
L	0.98***	1.01***	0.12***	0.56***	54.71%
	(26.78)	(36.80)	(2.73)	(11.61)	
S	-0.72***	1.11***	0.35***	0.61***	68.67%
	(-17.99)	(48.03)	(8.57)	(12.12)	
			Day 1		
L-S	0.04***	-0.11***	-0.24***	-0.20***	16.41%
	(2.84)	(-4.72)	(-6.44)	(-5.47)	
L	0.02**	0.97***	0.13***	0.34***	86.07%
	(2.30)	(67.58)	(6.31)	(12.52)	
S	-0.02*	1.08***	0.37***	0.54***	86.60%
	(-1.70)	(52.70)	(10.54)	(18.79)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 7: Returns on portfolios based on near and distant future news arriving at date 0

Panel A: Near future news ($offset < 60$)

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.15	5.57	0.03	1.09	0.12	7.03
-4	0.15	5.80	0.04	1.23	0.11	5.92
-3	0.15	5.56	0.02	0.59	0.13	7.19
-2	0.17	6.34	-0.01	-0.19	0.18	9.02
-1	0.50	13.43	-0.22	-6.07	0.72	19.99
0	1.47	25.66	-0.91	-19.19	2.37	30.57
1	0.08	2.91	0.02	0.76	0.05	2.76
2	0.07	2.46	0.05	1.53	0.02	1.04
3	0.05	2.13	0.05	1.66	0.00	0.05
4	0.06	2.28	0.07	2.12	-0.01	-0.56
5	0.06	2.49	0.05	1.47	0.01	0.84

Panel B: Distant future news ($offset \geq 60$)

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.12	4.61	0.02	0.77	0.09	5.38
-4	0.12	4.79	0.02	0.49	0.10	5.40
-3	0.13	5.11	0.03	0.89	0.10	5.41
-2	0.13	4.85	-0.01	-0.31	0.14	6.99
-1	0.30	10.63	-0.14	-4.14	0.44	17.92
0	0.56	17.51	-0.29	-8.62	0.86	29.39
1	0.09	3.60	0.04	1.25	0.05	2.97
2	0.07	2.80	0.06	1.97	0.01	0.38
3	0.05	2.03	0.05	1.59	0.00	0.20
4	0.06	2.38	0.06	1.99	-0.00	-0.08
5	0.05	2.10	0.03	1.07	0.02	1.25

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning near future news, i.e, offset < 60 (Panel A) and distant future news offset $\geq$ 60 (Panel B). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 8: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed based on near and distant future news

Panel A: Near future news ($offset \leq 60$)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	2.38***	-0.04	-0.17**	-0.03	0.39%
	(23.27)	(-0.71)	(-2.21)	(0.32)	
L	1.42***	1.04***	0.15**	0.64***	32.81%
	(21.72)	(23.45)	(2.15)	(7.64)	
S	-0.96***	1.08***	0.32***	0.61***	56.83%
	(-18.03)	(38.25)	(7.71)	(11.35)	
			Day 1		
L-S	0.05**	-0.10***	-0.21***	-0.14***	8.23%
	(2.58)	(-3.74)	(-4.87)	(-3.03)	
L	0.02	0.98***	0.13***	0.39***	79.39%
	(1.34)	(57.28)	(4.97)	(10.78)	
S	-0.03**	1.07***	0.33***	0.54***	82.98%
	(-2.55)	(51.50)	(9.40)	(15.58)	

(Continued on next page)

Table 8 – *(Continued from previous page)*

Panel B: Distant future news ($offset > 60$)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.85***	-0.09***	-0.26***	-0.11**	4.45%
	(27.24)	(-2.66)	(-4.69)	(-2.29)	
L	0.51***	1.01***	0.12***	0.43***	64.03%
	(23.07)	(49.17)	(3.57)	(12.77)	
S	-0.35***	1.10***	0.39***	0.54***	72.82%
	(-16.79)	(44.66)	(9.52)	(12.65)	
			Day 1		
L-S	0.05***	-0.11***	-0.23***	-0.19***	13.06%
	(2.86)	(-4.54)	(-6.39)	(-6.25)	
L	0.03***	0.98***	0.17***	0.28***	84.82%
	(2.65)	(67.47)	(7.81)	(12.99)	
S	-0.02	1.09***	0.40***	0.48***	83.66%
	(-1.31)	(42.91)	(10.50)	(15.82)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 9: Returns on portfolios based on news signals arriving at date 0, for news of different coverage

Panel A: News with only one mention on day 0

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.10	4.10	0.07	2.36	0.03	2.65
-4	0.10	4.47	0.06	2.28	0.04	3.57
-3	0.10	4.31	0.05	1.86	0.05	4.45
-2	0.12	4.92	0.03	1.21	0.08	6.76
-1	0.17	7.08	-0.06	-2.02	0.23	17.40
0	0.17	7.17	-0.05	-1.72	0.22	17.47
1	0.08	3.38	0.08	3.05	-0.01	-0.66
2	0.08	3.60	0.08	2.79	0.01	0.63
3	0.08	3.45	0.09	3.15	-0.01	-0.58
4	0.07	3.09	0.08	2.85	-0.01	-0.45
5	0.08	3.31	0.07	2.70	0.00	0.33

Panel B: News having two to five mentions on day 0

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.08	3.55	0.04	1.56	0.04	2.77
-4	0.12	4.96	0.05	1.68	0.07	3.99
-3	0.12	5.26	0.03	1.03	0.09	5.66
-2	0.12	5.17	-0.00	-0.11	0.13	7.91
-1	0.26	10.10	-0.14	-4.50	0.40	21.40
0	0.31	12.47	-0.22	-7.33	0.54	25.29
1	0.09	3.70	0.08	2.97	0.00	0.12
2	0.07	3.16	0.08	2.67	-0.01	-0.33
3	0.06	2.69	0.09	3.11	-0.02	-1.50
4	0.07	3.00	0.07	2.44	-0.00	-0.13
5	0.07	2.84	0.07	2.48	-0.01	-0.32

(Continued on next page)

Table 9 – *(Continued from previous page)*

Panel C: News having six to nine mentions on day 0

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.13	4.30	0.02	0.55	0.11	3.86
-4	0.12	4.20	0.03	0.93	0.09	3.18
-3	0.12	4.11	0.02	0.70	0.09	3.10
-2	0.14	4.83	-0.02	-0.63	0.17	5.01
-1	0.36	11.09	-0.23	-5.12	0.59	13.35
0	0.50	13.44	-0.54	-11.70	1.04	21.79
1	0.11	3.75	0.04	1.12	0.07	2.03
2	0.13	3.41	0.10	2.94	0.04	0.86
3	0.11	3.19	0.11	3.12	-0.00	-0.04
4	0.07	2.51	0.06	1.81	0.01	0.17
5	0.09	3.16	0.07	2.10	0.01	0.44

Panel D: News having more than ten mentions on day 0

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.11	3.23	0.01	0.12	0.10	2.14
-4	0.14	3.84	0.04	0.90	0.10	2.18
-3	0.16	4.91	-0.02	-0.52	0.18	4.23
-2	0.15	4.15	-0.11	-2.33	0.25	5.82
-1	0.36	8.76	-0.51	-9.52	0.87	14.62
0	0.97	13.77	-1.06	-14.61	2.03	20.19
1	0.11	2.79	0.01	0.22	0.09	2.06
2	0.07	2.30	0.10	2.42	-0.03	-0.75
3	0.10	3.18	0.15	2.86	-0.05	-0.98
4	0.05	1.75	0.17	2.91	-0.11	-2.01
5	0.07	2.40	0.11	1.87	-0.04	-0.64

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning news with one mention (Panel A), two to five mentions (Panel B), six to nine mentions (Panel C) and more than ten mentions (Panel D) of the same KPI on the same day. The result is for the sub-sample of stocks in the Russell 1000 index. We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 10: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed on high-coverage and low-coverage news

Panel A: News having only one mention on day 0					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.22***	-0.08***	-0.19***	-0.12***	14.32%
	(17.72)	(-4.15)	(-7.56)	(-5.49)	
L	0.11***	0.99***	0.11***	0.16***	92.35%
	(14.87)	(74.58)	(7.84)	(9.19)	
S	-0.11***	1.07***	0.30***	0.28***	88.78%
	(-9.70)	(61.34)	(12.30)	(11.65)	
			Day 1		
L-S	-0.02*	-0.04**	-0.18***	-0.15***	13.47%
	(-1.65)	(-2.46)	(-7.39)	(-5.49)	
L	0.01**	1.00***	0.12***	0.15***	93.65%
	(2.37)	(142.60)	(8.70)	(8.96)	
S	0.03***	1.03***	0.30***	0.30***	88.26%
	(2.82)	(72.87)	(11.16)	(9.55)	

(Continued on next page)

Table 10 – *(Continued from previous page)*

Panel B: News having two to five mention on day 0					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.54***	-0.10***	-0.20***	-0.09**	7.03%
	(21.81)	(-3.61)	(-4.49)	(-2.04)	
L	0.26***	0.99***	0.12***	0.16***	84.22%
	(19.99)	(79.56)	(7.06)	(7.90)	
S	-0.29***	1.10***	0.32***	0.25***	79.27%
	(-14.82)	(45.59)	(8.37)	(6.82)	
			Day 1		
L-S	-0.00	-0.09***	-0.28***	-0.15***	15.85%
	(-0.32)	(-3.44)	(-7.17)	(-3.37)	
L	0.03***	0.97***	0.07***	0.10***	89.22%
	(2.95)	(114.26)	(3.60)	(4.87)	
S	0.03**	1.06***	0.35***	0.25***	83.33%
	(2.06)	(45.85)	(10.40)	(7.39)	

(Continued on next page)

Table 10 – *(Continued from previous page)*

Panel C: News having six to nine mentions on day 0					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.05***	-0.20***	-0.30***	-0.14	2.97%
	(20.86)	(-2.90)	(-4.63)	(-1.46)	
L	0.45***	0.93***	0.06*	0.14***	38.59%
	(14.63)	(16.46)	(1.67)	(2.60)	
S	-0.60***	1.13***	0.36***	0.28***	38.55%
	(-15.49)	(32.05)	(6.11)	(3.88)	
			Day 1		
L-S	0.06**	-0.05	-0.29***	-0.25***	4.76%
	(1.97)	(-1.05)	(-4.33)	(-3.02)	
L	0.05***	1.00***	0.04	-0.00	62.87%
	(2.76)	(26.42)	(1.12)	(-0.04)	
S	-0.01	1.05***	0.34***	0.25***	55.72%
	(-0.49)	(32.96)	(6.36)	(3.96)	

(Continued on next page)

Table 10 – *(Continued from previous page)*

Panel D: News having more than ten mentions on day 0					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	2.02***	-0.07	-0.34**	-0.26	0.73%
	(19.14)	(-0.77)	(-2.10)	(-1.35)	
L	0.91***	1.01***	0.06	0.18**	13.55%
	(12.23)	(13.41)	(0.47)	(2.02)	
S	-1.11***	1.08***	0.41***	0.43***	17.13%
	(-15.39)	(21.12)	(4.16)	(2.40)	
			Day 1		
L-S	0.08*	-0.12**	-0.40***	-0.34***	5.16%
	(1.72)	(-2.01)	(-3.29)	(-3.84)	
L	0.02	0.96***	0.04	-0.06	39.66%
	(0.70)	(27.89)	(0.69)	(-1.06)	
S	-0.06*	1.08***	0.45***	0.28***	47.01%
	(-1.85)	(21.89)	(4.65)	(4.67)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 11: Returns on portfolios based on news signals arriving at date 0, fresh news versus stale news

Panel A: Fresh news

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.12	4.02	0.07	2.51	0.04	2.43
-4	0.12	4.16	0.08	2.63	0.04	2.09
-3	0.11	3.67	0.05	1.48	0.06	3.33
-2	0.13	4.24	0.09	2.57	0.04	2.01
-1	0.30	8.81	-0.02	-0.66	0.32	11.17
0	1.25	22.06	-0.53	-13.50	1.77	33.28
1	0.12	4.33	0.01	0.38	0.11	5.53
2	0.07	2.37	0.07	2.18	-0.00	-0.05
3	0.08	2.87	0.05	1.51	0.03	1.64
4	0.05	1.84	0.06	1.94	-0.01	-0.45
5	0.05	2.01	0.07	2.26	-0.01	-0.67

Panel B: Stale news

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.14	5.70	0.06	1.73	0.08	5.50
-4	0.13	5.35	0.06	1.90	0.07	4.42
-3	0.14	5.81	0.06	1.87	0.08	5.38
-2	0.18	7.24	0.01	0.46	0.16	9.97
-1	0.44	15.84	-0.36	-10.47	0.81	34.20
0	0.66	21.46	-0.54	-15.26	1.20	37.63
1	0.08	3.22	0.04	1.18	0.04	2.98
2	0.06	2.43	0.06	1.87	0.00	0.09
3	0.05	2.15	0.06	1.85	-0.00	-0.31
4	0.06	2.50	0.07	2.23	-0.01	-0.64
5	0.06	2.48	0.05	1.69	0.01	0.47

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning fresh news ($novelty \geq 0.5$) and stale news ($novelty < 0.5$). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 12: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed on fresh and stale news

Panel A: Fresh news					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.77***	-0.01	-0.04	0.01	0.02%
	(30.09)	(-0.27)	(-0.45)	(0.14)	
L	1.18***	1.09***	0.24***	0.64***	29.74%
	(22.24)	(24.98)	(3.64)	(8.78)	
S	-0.59***	1.11***	0.28***	0.63***	59.07%
	(-18.58)	(36.18)	(6.39)	(13.89)	
			Day 1		
L-S	0.08***	-0.02	-0.09**	-0.14***	2.09%
	(4.47)	(-0.80)	(-2.39)	(-3.04)	
L	0.05***	1.01***	0.22***	0.42***	77.74%
	(3.57)	(46.45)	(10.86)	(15.58)	
S	-0.03***	1.03***	0.31***	0.56***	80.07%
	(-2.30)	(55.03)	(9.16)	(13.96)	

(Continued on next page)

Table 12 – *(Continued from previous page)*

Panel B: Stale news					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.20***	-0.10***	-0.23***	-0.22***	6.84%
	(28.62)	(-3.08)	(-5.11)	(-4.17)	
L	0.60***	1.12***	0.09***	0.45***	77.82%
	(27.25)	(44.11)	(3.05)	(16.68)	
S	-0.60***	1.12***	0.32***	0.67***	77.90%
	(-21.27)	(52.74)	(8.45)	(15.89)	
			Day 1		
L-S	0.04***	-0.10***	-0.28***	-0.28***	26.58%
	(3.20)	(-6.07)	(-6.21)	(-6.41)	
L	0.02**	0.98***	0.07***	0.33***	91.44%
	(2.23)	(107.18)	(3.72)	(15.24)	
S	-0.02**	1.08***	0.35***	0.62***	89.91%
	(-2.16)	(70.19)	(10.49)	(20.65)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 13: Returns on portfolios constructed on signals of fundamental news

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.11	4.46	0.07	2.32	0.04	4.23
-4	0.11	4.43	0.07	2.50	0.04	4.17
-3	0.12	4.65	0.06	2.20	0.06	5.98
-2	0.14	5.30	0.04	1.47	0.09	9.02
-1	0.23	8.86	-0.08	-2.67	0.31	19.02
0	0.29	10.75	-0.16	-4.93	0.45	18.45
1	0.07	2.59	0.04	1.49	0.02	2.29
2	0.06	2.48	0.06	2.06	0.00	0.31
3	0.06	2.27	0.06	2.01	-0.00	-0.05
4	0.05	2.13	0.06	2.20	-0.01	-1.11
5	0.05	2.07	0.05	1.73	0.00	0.23

NOTE: On every trading day, we sorted all stocks on the basis of *ptp* constructed on the sample of all tenses and of all KPIs except for "stock price" mentioned in past and present tense. We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolios in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 14: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed on signals based on fundamental news

Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.45***	-0.07***	-0.16***	0.01	5.19%
	(14.68)	(-3.79)	(-6.99)	(0.15)	
L	0.24***	1.03***	0.17***	0.56***	93.81%
	(19.54)	(132.38)	(12.50)	(23.48)	
S	-0.22***	1.10***	0.32***	0.56***	84.74%
	(-9.65)	(61.23)	(11.94)	(16.47)	
			Day 1		
L-S	0.02*	-0.05***	-0.16***	-0.09***	16.53%
	(1.82)	(-5.17)	(-7.56)	(-3.71)	
L	0.01	1.01***	0.16***	0.50***	96.45%
	(0.96)	(125.73)	(13.47)	(37.58)	
S	-0.01	1.06***	0.31***	0.59***	92.32%
	(-1.36)	(90.80)	(11.89)	(18.81)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 15: Returns on portfolios based on news signals arriving at date 0, different types of news

Panel A: Financial news

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.14	0.54	0.06	2.00	0.07	5.13
-4	0.14	0.56	0.06	1.87	0.08	5.70
-3	0.14	5.64	0.06	1.87	0.08	5.99
-2	0.17	6.87	0.02	0.63	0.15	10.39
-1	0.43	15.31	-0.34	-10.16	0.77	34.22
0	0.75	24.08	-0.55	-15.80	1.30	43.09
1	0.08	3.20	0.03	1.13	0.04	3.16
2	0.06	2.39	0.06	1.93	-0.00	-0.06
3	0.06	2.28	0.06	1.84	0.00	0.00
4	0.06	2.52	0.06	2.09	-0.00	-0.29
5	0.05	2.22	0.05	1.68	0.00	0.11

Panel B: ESG news, Long portfolio (ptp $\geq 90th - percencile$), short portfolio (ptp $< 10th - percencile$)

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.06	2.25	0.08	2.98	-0.02	-1.23
-4	0.08	3.17	0.08	3.07	-0.01	-0.27
-3	0.05	2.09	0.07	2.63	-0.02	-1.08
-2	0.07	2.59	0.06	2.38	0.00	0.16
-1	0.08	3.29	0.07	2.30	0.03	0.67
0	0.09	3.18	0.09	2.98	0.02	-0.26
1	0.06	2.35	0.06	2.31	-0.01	-0.24
2	0.08	2.55	0.06	2.35	0.01	0.62
3	0.06	2.53	0.08	2.43	-0.02	-0.53
4	0.03	1.32	0.08	2.67	-0.05	-2.19
5	0.07	2.76	0.06	2.39	0.01	0.51

(Continued on next page)

Table 15 – *(Continued from previous page)*

Panel C: ESG news, Long portfolio (ptp $\geq$ 0.5), short portfolio (ptp $<$ 0.5)						
	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.07	3.07	0.07	2.88	0.00	0.06
-4	0.07	2.95	0.09	3.27	-0.02	-1.41
-3	0.07	2.99	0.07	2.69	-0.00	-0.06
-2	0.07	2.78	0.06	2.13	0.01	0.66
-1	0.08	3.53	0.06	2.01	0.02	1.16
0	0.09	3.86	0.08	2.74	0.01	0.51
1	0.07	3.00	0.07	2.57	0.00	0.18
2	0.06	2.53	0.07	2.53	-0.01	-0.46
3	0.05	2.32	0.07	2.37	-0.02	-0.90
4	0.06	2.41	0.09	2.89	-0.03	-1.61
5	0.06	2.47	0.07	2.61	-0.01	-0.69

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning financial KPIs (Panel A) and ESG KPIs (Panel B and C). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile in Panel A and B. In Panel C, we took a long position in stocks with *ptp* higher than 0.5 and a short position in those with *ptp* lower than 0.5. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 16: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed based on news signals, different types of news

Panel A: Financial news					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.29***	-0.08***	-0.21***	-0.18***	5.50%
	(33.41)	(-3.16)	(-4.94)	(-3.50)	
L	0.70***	1.02***	0.10***	0.48***	75.73%
	(31.44)	(61.94)	(3.47)	(15.89)	
S	-0.61***	1.10***	0.31***	0.67***	79.22%
	(-22.71)	(56.37)	(9.00)	(17.54)	
			Day 1		
L-S	0.04***	-0.09***	-0.27***	-0.24***	25.2%
	(3.39)	(-5.75)	(-6.19)	(-6.19)	
L	0.01**	0.99***	0.08***	0.37***	92.86%
	(2.04)	(108.28)	(4.78)	(23.10)	
S	-0.03**	1.08***	0.34***	0.61***	
	(-2.50)	(114.73)	(27.33)	(35.98)	

(Continued on next page)

Table 16 – *(Continued from previous page)*

Panel B: ESG news, Long portfolio (ptp $\geq 90th - percencile$), short portfolio (ptp $< 10th - percencile$)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	-0.01	-0.07***	0.05	-0.07	0.74%
	(-0.40)	(-2.72)	(1.04)	(-1.41)	
L	0.03**	0.95***	0.27***	0.17***	70.33%
	(2.07)	(64.91)	(9.96)	(5.95)	
S	0.04**	1.02***	0.22***	0.24***	62.81%
	(2.04)	(47.21)	(5.52)	(6.31)	
			Day 1		
L-S	-0.01	0.02	0.03	-0.15***	0.67%
	(-0.31)	(0.68)	(0.77)	(-2.97)	
L	-0.00	1.02***	0.29***	0.23***	70.05%
	(-0.10)	(26.09)	(8.58)	(6.21)	
S	0.00	1.00	0.26***	0.38***	62.25%
	(0.18)	(30.69)	(6.17)	(9.18)	

(Continued on next page)

Table 16 – *(Continued from previous page)*

Panel C: ESG news, Long portfolio (ptp $\geq$ 0.5), short portfolio (ptp $<$ 0.5)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.01	-0.03	-0.00	0.04	0.16%
	(0.29)	(-1.60)	(-0.04)	(0.95)	
L	0.04***	1.00***	0.21***	0.25***	91.00%
	(4.48)	(102.18)	(14.62)	(11.42)	
S	0.03	1.02***	0.21***	0.21***	64.37%
	(1.60)	(53.39)	(5.80)	(5.92)	
			Day 1		
L-S	0.00	-0.03	-0.00	-0.03	0.48%
	(0.21)	(-1.55)	(-0.06)	(-1.22)	
L	0.01*	0.98***	0.21***	0.22***	92.09%
	(1.75)	(125.45)	(12.52)	(11.13)	
S	0.01	1.01***	0.21***	0.25***	79.78%
	(0.60)	(49.73)	(8.23)	(9.71)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table 17: Returns on portfolios based on news signals arriving at date 0, different firm sizes

Panel A: Russell 1000 components

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.11	4.94	0.06	1.92	0.06	4.08
-4	0.12	5.08	0.05	1.79	0.06	4.50
-3	0.12	5.05	0.03	0.92	0.09	6.58
-2	0.14	5.96	0.00	0.03	0.14	9.04
-1	0.27	11.08	-0.17	-5.65	0.44	26.66
0	0.41	16.18	-0.26	-8.42	0.67	32.05
1	0.09	3.86	0.07	2.55	0.01	0.76
2	0.07	3.13	0.08	2.65	-0.00	-0.26
3	0.07	2.86	0.08	2.98	-0.02	-1.17
4	0.07	3.02	0.08	2.90	-0.02	-1.02
5	0.07	3.04	0.08	2.78	-0.01	-0.86

Panel B: Russell 1000 components excluded

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.15	5.31	0.08	2.21	0.07	4.20
-4	0.15	5.33	0.07	2.06	0.08	4.55
-3	0.16	5.52	0.08	2.41	0.08	4.16
-2	0.20	6.88	0.04	1.19	0.16	8.17
-1	0.56	16.01	-0.46	-12.08	1.02	32.38
0	1.07	25.07	-0.71	-18.99	1.79	43.90
1	0.07	2.60	0.00	0.06	0.07	4.25
2	0.04	1.53	0.04	1.22	0.00	0.10
3	0.05	1.68	0.04	1.10	0.10	0.56
4	0.05	1.78	0.05	1.59	-0.00	-0.27
5	0.04	1.33	0.04	1.14	-0.00	-0.07

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples of stocks that belong to the Russell 1000 index (Panel A) and stocks that do not belong to the Russell 1000 index (Panel B). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table 18: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios based on news signals, different firm sizes

Panel A: Russell 1000					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	0.67***	-0.11***	-0.26***	-0.13***	12.72%
	(28.31)	(-3.81)	(-5.99)	(-3.05)	
L	0.35***	0.99***	0.06***	0.17***	86.14%
	(27.29)	(68.23)	(4.14)	(9.73)	
S	-0.32***	1.11***	0.32***	0.30***	83.60%
	(-18.79)	(46.20)	(7.98)	(8.91)	
			Day 1		
L-S	-0.00	-0.08***	-0.30***	-0.18***	20.32%
	(-0.13)	(-4.69)	(-6.96)	(-5.10)	
L	0.02***	0.98***	0.05***	0.12***	92.93%
	(3.25)	(130.29)	(3.42)	(8.74)	
S	0.02*	1.06***	0.35***	0.29***	84.74%
	(1.85)	(64.92)	(9.39)	(9.18)	

(Continued on next page)

Table 18 – *(Continued from previous page)*

Panel B: Russell 1000 excluded					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day 0		
L-S	1.79***	-0.04	-0.14**	-0.12*	1.07%
	(33.71)	(-1.24)	(-2.41)	(-1.66)	
L	1.02***	1.05***	1.16***	0.79***	54.74%
	(26.53)	(42.68)	(3.21)	(15.37)	
S	-0.77***	1.09***	0.30***	0.91***	87.35%
	(-24.58)	(49.73)	(8.30)	(17.98)	
			Day 1		
L-S	0.07***	-0.07***	-0.20***	-0.21***	10.52%
	(4.44)	(-4.22)	(-4.49)	(-4.49)	
L	0.01	1.01***	0.15***	0.62***	84.97%
	(0.86)	(60.67)	(6.22)	(26.66)	
S	-0.06***	1.08***	0.35***	0.83***	87.05%
	(-4.71)	(92.55)	(22.37)	(39.43)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

References

R. M. Anderson, K. S. Eom, S. B. Hahn, and J.-H. Park. Sources of stock return autocorrelation. In *Working paper*.

A. Aouadi and S. Marsat. Do ESG controversies matter for firm value? evidence from international data. *Journal of Business Ethics*, 151(4): 1027–1047, 01 Sep 2018.

B. M. Barber and T. Odean. 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, 01 Apr 2008.

N. Barberis, R. Greenwood, L. Jin, and A. Shleifer. X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115 (1):1–24, 01 Jan 2015.

J. R. Bence. Analysis of short time series: correcting for autocorrelation. *Ecology*, 76(2):628–639, 01 Mar 1995.

J. Boudoukh, R. Feldman, S. Kogan, and M. Richardson. Information, trading, and volatility: Evidence from firm-specific news. *The Review of Financial Studies*, 32(3):992–1033, 2019.

J. Y. Campbell. A variance decomposition for stock returns. *NBER working paper*, Jan 1990.

G. Capelle-Blancard and A. Petit. Every little helps? ESG news and stock market reaction. *Journal of Business Ethics*, 157(2):543–565, 30 Jun 2019.

T. Change and F. Change. Do stock prices move too much to be justified by subsequent changes in dividends?: Comment. *The American Economic Review*, 73(1):234–235, Mar 1983.

J. Conrad and G. Kaul. Time-variation in expected returns. *Journal of business*, pages 409–425, 1988.

J. Conrad and G. Kaul. Mean reversion in short-horizon expected returns. *The Review of Financial Studies*, 2(2):225–240, 1989.

B. Cui and P. Docherty. Stock price overreaction to ESG controversies. *Available at SSRN 3559915*, 17 Mar 2020.

D. M. Cutler, J. M. Poterba, and L. H. Summers. What moves stock prices? *NBER working paper*, Mar 1988.

A. K. Davis, J. M. Piger, L. M. Sedor, et al. *Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases*, volume 5. Citeseer, Jan 2006.

J. B. De Long, A. Shleifer, L. H. Summers, and R. J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *the Journal of Finance*, 45(2):379–395, Jun 1990.

E. Demers and C. Vega. Soft information in earnings announcements: News or noise? *International Finance Discussion Papers*, (951), Oct 2008.

F. Derrien, P. Krueger, A. Landier, and T. Yao. How do ESG incidents affect firm value? *Available at SSRN 3903274*, 2021.

E. F. Fama and K. R. French. Dissecting anomalies. *The Journal of Finance*, 63(4):1653–1678, 2008.

S. L. Heston and N. R. Sinha. News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal*, 73(3):67–83, 01 Jul 2017.

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

K. Hou, C. Xue, and L. Zhang. Replicating anomalies. *The Review of Financial Studies*, 33(5):2019–2133, 01 May 2020.

N. Jegadeesh and D. Wu. Word power: A new approach for content analysis. *Journal of Financial Economics*, 110(3):712–729, 01 Dec 2013.

Z. T. Ke, B. T. Kelly, and D. Xiu. Predicting returns with text data. Technical report, National Bureau of Economic Research, 02 Sep 2019.

P. Krüger. Corporate goodness and shareholder wealth. *Journal of Financial Economics*, 115(2):304–329, 01 Feb 2015.

A. W. Lo and A. C. MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1):41–66, 1988.

T. Loughran and B. McDonald. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65, Feb 2011.

G. Serafeim and A. Yoon. Stock price reactions to ESG news: The role of ESG ratings and disagreement. *Review of Accounting Studies*, pages 1–31, 10 Mar 2022.

P. C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 08 May 2007.

P. C. Tetlock, M. Saar-Tsechansky, and S. Macskassy. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467, Jun 2008.

5 APPENDIX

5.1 Correction for auto-correlation in the series of daily returns

Bence (1995) introduces a correction factor to the standard error of the mean for an order of sequence $y_i = \mu_i + \epsilon_i$ for $i = 1, 2, ..., T$ where T is the length of the series. In our case, $\mu_i = \mu$ as there is no explanatory variable in the model.

Let $\hat{\mu}$ be an estimate of μ, the residual $\hat{r_i}$ is obtained as $y_i - \hat{\mu}$. Assume the error terms follow an AR(1) process: $\epsilon_t = \rho\epsilon_{t-1} + a_t$.

The estimate of the sample mean is the sample average $\hat{\mu} = \frac{\sum_{i=1}^{T} y_i}{T} \equiv \overline{y}$. As the error terms are not i.i.d., the usual estimator of the standard error:

$$s = \sqrt{\frac{\sum_{i=1}^{T}(y_i - \hat{\mu})^2}{T(T-1)}} \tag{2}$$

would be downward biased by a factor k:

$$k = \left[\frac{1 + 2\delta/T}{1 - 2\delta/T(T-1)}\right]^{1/2} \tag{3}$$

where:

$$\delta = \frac{[(T-1)\rho - T\rho^2 + \rho^{T+1}]}{(1-\rho)^2} \tag{4}$$

k is a the correction factor such that $E(ks) = \sigma_{\overline{y}}$, where $\sigma_{\overline{y}}$ is the true standard deviation of the sample average $\overline{y}$.

First, we obtain the usual naïve standard error as in 2.

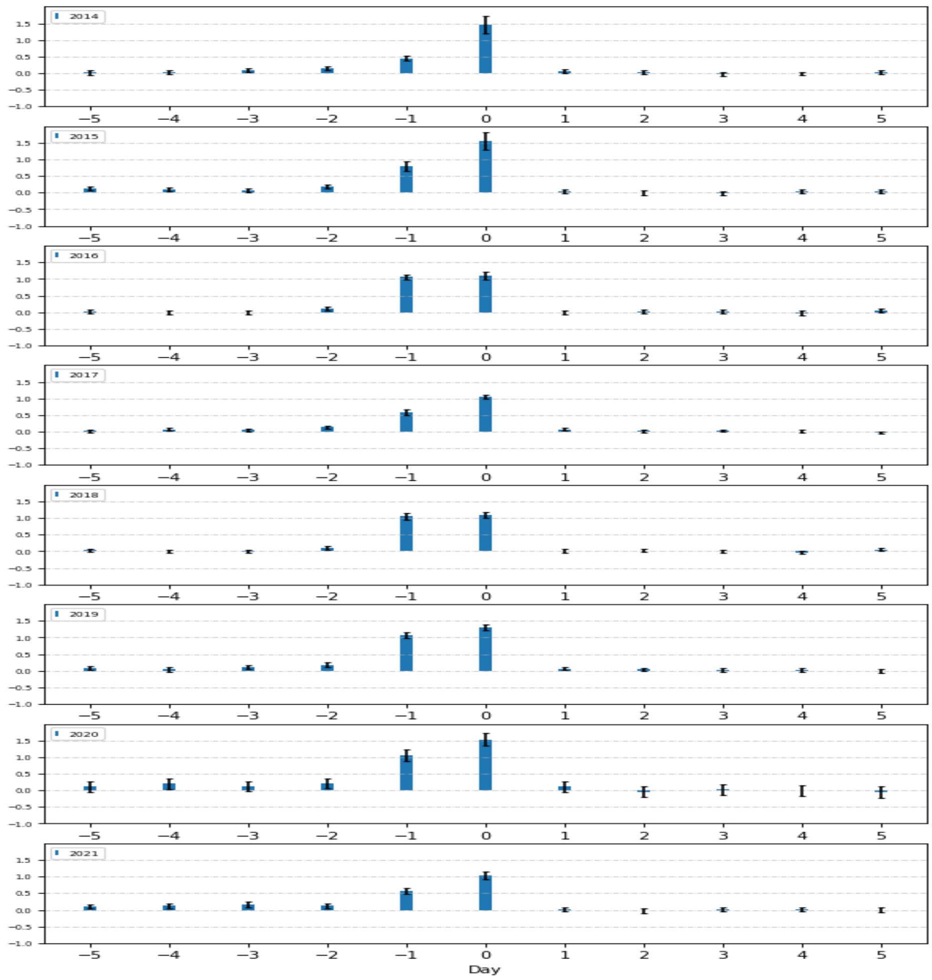
Second, we use OLS to obtain $\hat{\rho}$, an estimate of ρ. This allows us to compute the correction factor k.

Finally, we apply the correction factor k to the naïve standard error s to obtain $\tilde{s} = ks$.

In our data, all the estimates of $\hat{\rho}$ across various sample are mild (less than 0.2), therefore, all the factors k just barely exceed 1.

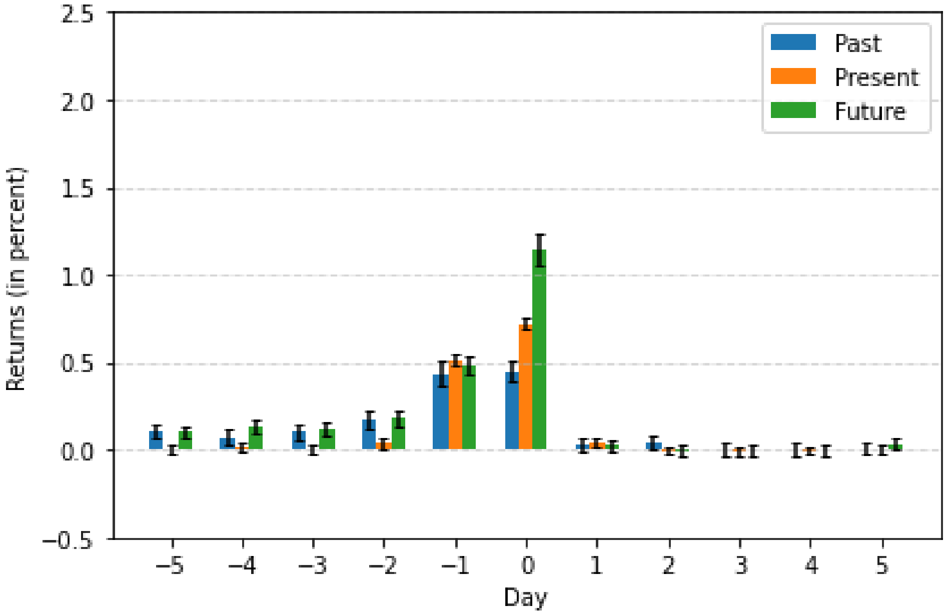
5.2 Figures and tables

Figure A1: Return on the long-short portfolio constructed on news signals by year



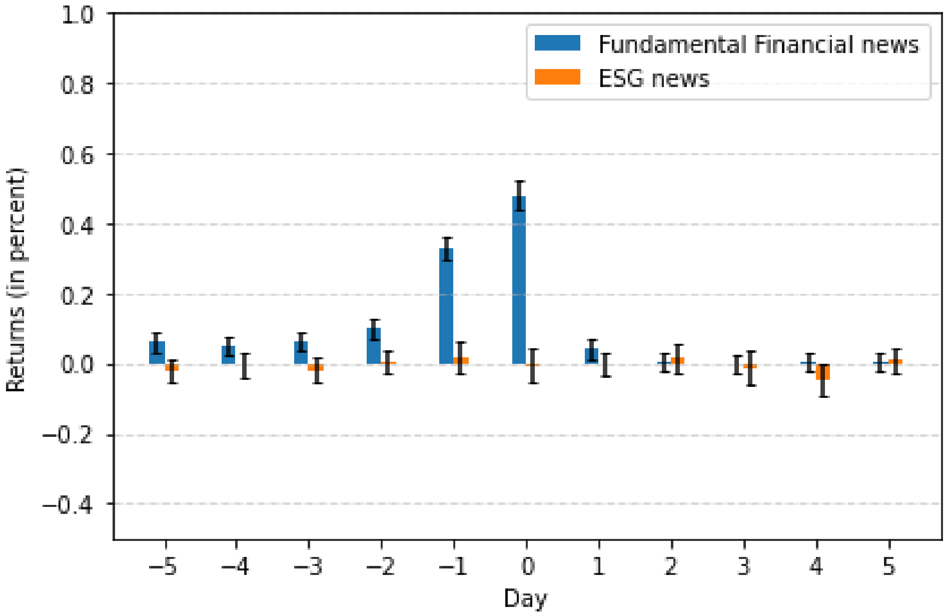
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-sample sample by year. The y-axis is in percentage points. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure A2: Return on the Portfolios constructed on news of different tenses



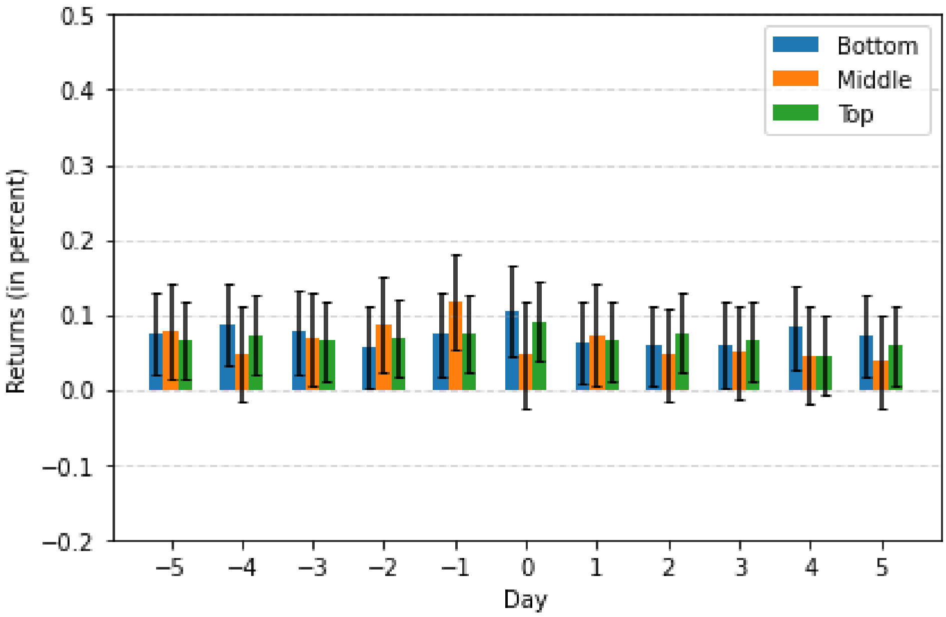
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples based on tenses. We make sure that stocks that appears in portfolios of one tense do not appear in other portfolios of other tenses, i.e, stocks that appear in the long or short portfolio for past news do not appear in any portfolios of present or future news, etc.. The error bars are the 95% confidence interval.

Figure A3: Return on the portfolios for fundamental financial news and ESG news



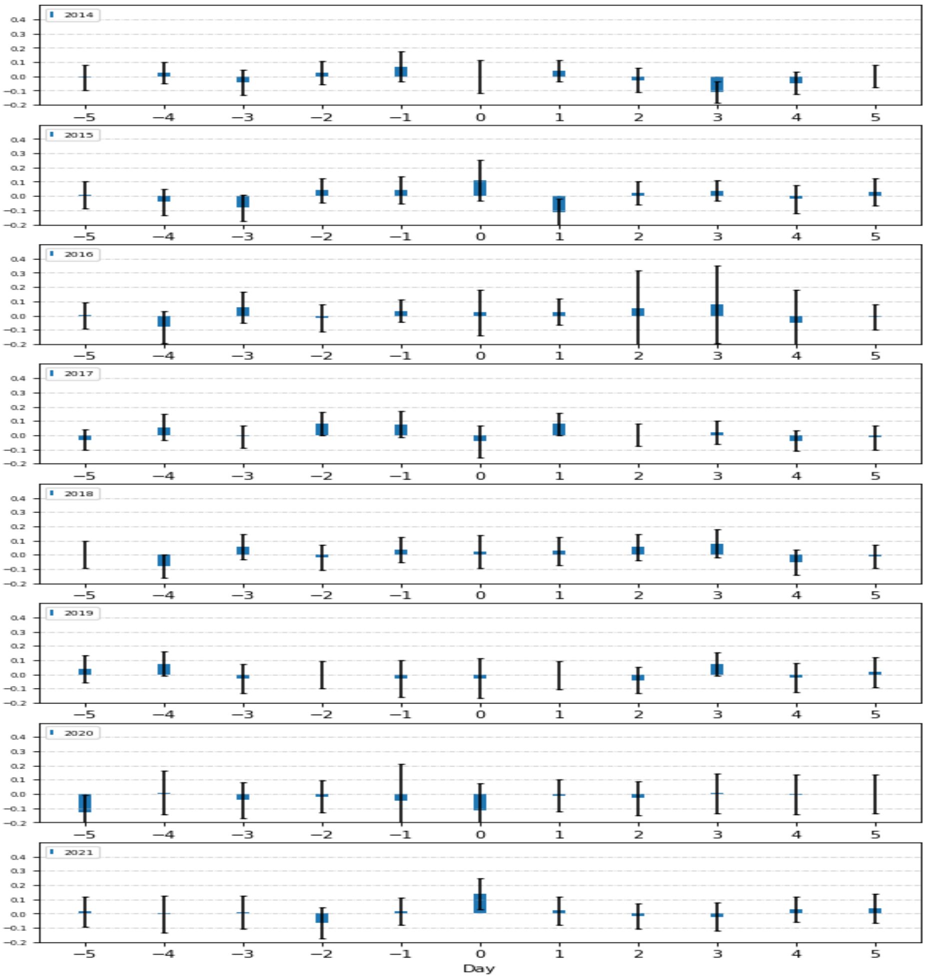
The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples concerning fundamental financial news and ESG news. For comparison purpose, the ESG portfolios are constructed in the same way as the fundamental financial portfolios (with stocks whose ptp are higher than the 90^{th} percentile making up the long portfolio and those whose ptp are lower than the 10^{th} percentile making up the short portfolio). The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Figure A4: Return on the Portfolios constructed on ESG news



The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the ESG sub-sample. The bottom portfolio consists of stocks that have *ptp* less than or equal to the 25%-quantile in the same day. The top portfolio consists of stocks that have *ptp* higher or equal to the 75%-quantile in the same day. The middle portfolio consists of stocks that have *ptp* belonging to the inter-quartile of the same day. The y-axis is in percentage points. The error bars are the 95% confidence interval.

Figure A5: Return on the long-short portfolio constructed on ESG signals by year



The bar charts present the average returns of the Long - Short strategy for the period [-5, +5] days around the portfolio construction days, on the sub-samples by year. The y-axis is in percentage points. The error bars are the 95% confidence interval computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table A1: Returns on portfolios based on news signals on date 0, different tenses, excluding stocks appearing in different tense portfolios in the same day

Panel A: News about past events

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.13	4.48	0.03	0.90	0.11	4.86
-4	0.13	4.28	0.06	1.60	0.07	3.28
-3	0.13	4.56	0.03	0.71	0.10	4.61
-2	0.21	6.21	0.04	0.99	0.17	6.32
-1	0.30	9.17	-0.13	-3.21	0.43	12.35
0	0.42	11.26	-0.03	-0.96	0.45	15.40
1	0.10	3.24	0.07	2.07	0.03	1.22
2	0.08	3.06	0.04	1.34	0.04	1.88
3	0.05	1.96	0.05	1.54	0.00	0.07
4	0.07	2.28	0.06	1.81	0.00	0.02
5	0.06	2.24	0.05	1.61	0.01	0.46

Panel B: News about present events

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.10	4.18	0.10	3.26	0.01	0.54
-4	0.12	4.68	0.10	3.35	0.02	1.52
-3	0.12	4.63	0.12	4.08	-0.00	-0.01
-2	0.13	5.10	0.09	3.06	0.04	2.51
-1	0.30	11.20	-0.21	-6.94	0.51	28.83
0	0.43	14.96	-0.30	-10.21	0.72	38.63
1	0.08	3.07	0.04	1.37	0.04	3.04
2	0.05	2.17	0.06	2.06	-0.01	-0.46
3	0.06	2.46	0.07	2.63	-0.01	-1.12
4	0.05	2.11	0.05	1.93	-0.00	-0.34
5	0.05	2.17	0.05	1.85	0.00	0.12

(Continued on next page)

Table A1 – *(Continued from previous page)*

Panel C: News about future events						
	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.12	4.76	0.02	0.74	0.10	5.75
-4	0.15	5.77	0.02	0.71	0.13	6.63
-3	0.15	5.54	0.03	0.79	0.12	6.56
-2	0.17	6.50	-0.01	-0.25	0.18	8.31
-1	0.37	10.78	-0.11	-3.26	0.48	17.35
0	0.83	21.26	-0.31	-9.05	1.14	26.47
1	0.07	2.48	0.04	1.39	0.02	1.29
2	0.05	1.96	0.05	1.77	-0.00	-0.21
3	0.05	2.01	0.06	1.83	-0.00	-0.08
4	0.05	1.83	0.05	1.60	-0.00	-0.19
5	0.06	2.42	0.03	1.07	0.03	1.86

NOTE: On every trading day, we sorted all stocks based on *ptp* constructed on the sub-samples concerning past news (Panel A), present news (Panel B) and future news (Panel C). We took a long position in stocks whose *ptp* signals are equal or higher than the 90^{th} percentile and a short position in stocks whose *ptp* signals are equal or lower than the 10^{th} percentile. The table reports the average returns (in percentage points) of such portfolio in the period [-5, +5] days around the portfolio construction days. The difference between this table and table 5 is that we do not allow stocks to appear in different tense portfolios in the same day. For example, stocks that appear in the past-tense (long or short) portfolio do not appear in any present-tense or future-tense portfolio constructed in the same day. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table A2: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed based on news signals, different tenses, excluding stocks appearing in different tense portfolios in the same day

Panel A: News about past events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day-0		
L-S	0.45***	-0.02	-0.30***	-0.11***	6.33%
	(14.30)	(-0.83)	(-7.46)	(-1.92)	
L	0.36***	1.02***	0.06**	0.44***	67.34%
	(14.11)	(38.65)	(2.35)	(8.90)	
S	-0.10***	1.04***	0.36***	0.55***	77.26%
	(-5.23)	(51.88)	(8.85)	(15.90)	
			Day-1		
L-S	0.02	-0.04	-0.31***	-0.10***	11.14%
	(1.14)	(-1.47)	(-6.88)	(-2.14)	
L	0.03**	1.05***	0.09***	0.45***	82.53%
	(2.14)	(45.14)	(3.99)	(10.63)	
S	0.01	1.09***	0.40***	0.55***	82.32%
	(0.40)	(54.84)	(9.73)	(19.63)	

(Continued on next page)

Table A2 – *(Continued from previous page)*

Panel B: News about present events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day-0		
L-S	0.72***	-0.07***	-0.11***	-0.07***	3.13%
	(34.09)	(-2.49)	(-3.53)	(-1.71)	
L	0.37***	1.01***	0.17***	0.47***	80.40%
	(23.33)	(57.55)	(8.63)	(14.57)	
S	-0.36***	1.08***	0.27***	0.53***	88.11%
	(-28.61)	(61.08)	(11.21)	(25.60)	
			Day-1		
L-S	0.04***	-0.04***	-0.15***	-0.20***	11.48%
	(3.09)	(-3.56)	(-5.25)	(-8.07)	
L	0.02*	1.00***	0.12***	0.36***	88.14%
	(1.79)	(178.35)	(7.69)	(24.95)	
S	-0.02***	1.05***	0.27***	0.56***	91.18%
	(-2.64)	(84.62)	(11.54)	(23.73)	

(Continued on next page)

Table A2 – *(Continued from previous page)*

Panel C: News about future events					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day-0		
L-S	1.15***	-0.10*	-0.16***	0.00	1.53%
	(19.91)	(-1.90)	(-3.29)	(0.17)	
L	0.78***	0.97***	0.19***	0.57***	0.78%
	(19.87)	(25.47)	(4.02)	(15.57)	
S	-0.37***	1.07***	0.35***	0.57***	72.99%
	(-13.08)	(42.34)	(9.90)	(11.03)	
			Day-1		
L-S	0.02	-0.07**	-0.19***	-0.21***	8.62%
	(0.91)	(-2.52)	(-5.68)	(-4.07)	
L	0.01	0.99***	0.15***	0.36***	80.81%
	(0.40)	(55.31)	(6.13)	(11.10)	
S	-0.01	1.06***	0.35***	0.57***	84.26%
	(-0.10)	(48.87)	(11.63)	(17.10)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. The difference between this table and table 6 is that we do not allow stocks to appear in different tense portfolios in the same day. For example, stocks that appear in the past-tense (long or short) portfolio do not appear in any present-tense or future-tense portfolio constructed in the same day. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.

Table A3: Top 20 KPIs mentioned near and distant future news

	Near future news (offset ≤ 60)			**Distant future news (offset > 60)**		
Rank	**KPI**	**Pct.**	**Cum. pct.**	**KPI**	**Pct.**	**Cum. pct.**
1	stock_price	23.06	23.06	revenue	23.01	23.01
2	revenue	13.84	36.90	profit	19.16	42.17
3	profit	13.02	49.91	stock_price	7.39	49.56
4	demand	1.70	54.61	production	5.02	54.58
5	production	4.23	58.84	demand	4.48	59.07
6	prices	3.97	62.82	prices	3.61	62.68
7	risk	3.96	66.77	expenses	3.16	65.84
8	employment	3.44	70.21	employment	3.10	68.94
9	expenses	2.65	72.87	investment	2.63	71.57
10	capital_return	2.43	75.30	risk	2.51	74.08
11	investment	2.18	77.48	capital_return	2.74	76.55
12	assets	1.95	79.44	market	2.20	78.75
13	analyst_ratings	1.89	81.33	assets	1.96	80.71
14	operations	1.86	83.18	operating expenses	1.74	82.45
15	market	1.75	84.93	operations	1.71	84.16
16	opportunity	1.42	86.35	analyst_ratings	1.45	85.60
17	operating_expenses	1.29	87.64	global_climate_change	1.34	86.94
18	future_stock_values	1.07	88.71	opportunities	1.22	88.17
19	valuation	0.98	89.69	cash_flow	1.14	89.30
20	taxes	0.77	90.46	liabilities	1.02	90.32

NOTE: The table reports the top 20 KPIs that appear in near and distant future news during the sample period. The second and fifth columns report the percentage of news statements that contains the given KPI. The third and sixth column report the cumulative percentage.

Table A4: Returns on portfolios based on ESG news signals on date 0

Panel A: Long portfolio (ptp ≥ 0.5), short portfolio (ptp < 0.5)

	Long portfolio		Short portfolio		Long-Short portfolio	
Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.05	2.06	0.07	2.58	-0.02	-0.78
-4	0.08	3.15	0.08	2.86	-0.00	-0.03
-3	0.04	1.42	0.08	2.80	-0.04	-2.02
-2	0.07	2.38	0.07	2.58	-0.01	-0.27
-1	0.08	2.98	0.07	2.40	0.01	0.57
0	0.08	2.71	0.07	2.42	0.00	0.06
1	0.06	2.23	0.05	1.75	0.01	0.47
2	0.07	2.02	0.05	1.97	0.02	0.60
3	0.05	1.98	0.08	2.51	-0.03	-1.18
4	0.03	0.97	0.07	2.61	-0.04	-2.20
5	0.08	3.00	0.07	2.72	0.01	0.51

Panel B: Long portfolio (ptp $\geq 90^{th} - percentile$), short portfolio (ptp $\leq 10^{th} - percentile$)

Date	**Mean**	**t-stat**	**Mean**	**t-stat**	**Mean**	**t-stat**
-5	0.07	3.00	0.07	2.64	0.00	0.08
-4	0.07	2.87	0.09	3.24	-0.02	-1.27
-3	0.07	2.85	0.08	2.76	-0.01	-0.41
-2	0.07	2.73	0.07	2.42	0.00	0.22
-1	0.08	3.58	0.06	2.09	0.02	1.47
0	0.09	3.56	0.06	2.06	0.02	1.24
1	0.06	2.60	0.05	1.78	0.01	0.84
2	0.06	2.38	0.06	2.24	-0.00	-0.16
3	0.05	2.16	0.08	2.31	-0.02	-0.04
4	0.05	2.26	0.08	2.84	-0.02	-1.44
5	0.06	2.72	0.07	2.85	-0.01	-0.61

NOTE: On every trading day, we sorted all stocks on the basis of *ptp* constructed on the ESG news sample. For the long portfolio portfolio, we choose stocks whose *ptp* signals are equal or higher than 0.5 (Panel A) and equal or higher than the 90^{th} percentile (Panel B). For the short portfolio, we choose stocks whose *ptp* signals are lower than 0.5 (Panel A) and lower than the 10^{th} percentile (Panel B). Stocks that appear in the ESG (long or short) portfolio do not appear in portfolios constructed with other KPIs in the same day. The table reports the average returns (in percentage points) of such portfolios in the period [-5, +5] days around the portfolio construction days. The t-statistics are computed using standard errors adjusted for auto-correlation in portfolio returns as in Bence (1995).

Table A5: Abnormal returns on day 0 and day 1 relative to the Fama-French three-factor model of portfolios constructed based on ESG news signals

Panel A: Long portfolio (ptp $\geq$ 0.5), short portfolio (ptp $<$ 0.5)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day-0		
L-S	-0.00	-0.04	0.03	-0.04	0.31%
	(-0.13)	(-1.64)	(0.55)	(-0.71)	
L	0.02	0.97***	0.27***	0.18***	71.17%
	(1.32)	(50.98)	(9.08)	(5.53)	
S	0.02	1.02***	0.24***	0.22***	63.97%
	(1.18)	(36.30)	(5.60)	(4.70)	
			Day-1		
L-S	0.01	0.03	0.05	-0.15***	0.57%
	(0.42)	(0.74)	(0.95)	(-3.03)	
L	-0.00	1.04***	0.32***	0.24***	63.58%
	(-0.04)	(27.12)	(9.05)	(6.03)	
S	-0.02	1.01***	0.27***	0.39***	58.98%
	(-0.72)	(23.71)	(4.92)	(8.61)	

(Continued on next page)

Table A5 – *(Continued from previous page)*

Panel B: Long portfolio (ptp $\geq 90^{th} - percentile$), short portfolio (ptp $\leq 10^{th} - percentile$)					
Formation	α	β_{mkt}	β_{hml}	β_{smb}	R^2
			Day-0		
L-S	0.02	-0.02	-0.01	0.07	0.25%
	(1.06)	(-0.91)	(-0.17)	(1.17)	
L	0.03***	1.00***	0.22***	0.27***	90.34%
	(3.72)	(101.39)	(13.85)	(9.01)	
S	0.01	1.02***	0.23***	0.20***	64.24%
	(0.56)	(39.03)	(5.67)	(4.22)	
			Day-1		
L-S	0.01	-0.02	0.01	-0.01	0.16%
	(0.73)	(-0.95)	(0.42)	(-0.15)	
L	0.00	0.99***	0.22***	0.24***	91.66%
	(0.60)	(104.41)	(13.17)	(10.37)	
S	-0.01	1.01***	0.21***	0.24***	76.16%
	(-0.66)	(40.61)	(7.08)	(8.36)	

The table reports the day 0 and day 1 alphas, the betas on market excess return (β_{mkt}), High Minus Low (β_{hml}) and Small Minus Big (β_{smb}) factors of the Fama-French three-factor models and the corresponding R-squared. For the long portfolio portfolio, we choose stocks whose *ptp* signals are equal or higher than 0.5 (Panel A) and equal or higher than the 90^{th} percentile (Panel B). For the short portfolio, we choose stocks whose *ptp* signals are lower than 0.5 (Panel A) and lower than the 10^{th} percentile (Panel B). Stocks that appear in the ESG (long or short) portfolio do not appear in portfolios constructed with other KPIs in the same day. T-statistics are computed using Newey-West (1987) standard errors that are robust to heteroskedasticity and auto-correlation up to eight lags.