Narrative Attention Pricing

Hojoon Lee Xiaoxia Lou Gideon Ozik Ronnie Sadka*

First draft: July 1, 2023

This version: October 12, 2023

Abstract

This paper demonstrates that economic narratives significantly price the cross-section of stocks. Using a vast dataset of more than 150k digital media sources since 2013, roughly 350 narratives are quantified, and corresponding narrative-mimicking, long-short portfolios are constructed using stock return narrative betas. Narrative-mimicking portfolios of recently trending narratives outperform those of descending attention by about 7% annually, controlling for standard risk factors. The cross-sectional narrative-beta-pricing is independent of past return and is neither significantly impacted by narrative coverage at the stock level nor earnings announcements. The results suggest that while investors respond to short-run narrative shocks as measured by narrative betas, they under-react to long-run narrative trends, manifesting narrative momentum returns.

*Lee: Carroll School of Management, Boston College, email: hojoon.lee@bc.edu. Lou: Alfred Lerner College of Business and Economics, University of Delaware, email: lous@udel.edu. Ozik: EDHEC Business School, email: gideon.ozik@edhec.com. Sadka: Carroll School of Management, Boston College, email: ronnie.sadka@bc.edu. Ozik, and Sadka are affiliated with MKT MediaStats, LLC (www.mktmediastats.com). The views expressed are solely of the authors. We thank Robert Korajczyk, Robert Macrae, Stephane Mattatia as well as participants at Chicago Quantitative Alliance conference Spring 2023, FOME Forum (Nice), State Street Research Retreat 2023, participants at 15th Annual Hedge Fund Research Conference for insightful comments.

1 Introduction

In his most recent seminal work, Shiller (2019) postulates that economic narratives drive markets over long horizons (years). He provides a few examples using low-frequency measures to gauge whether a narrative "goes viral," and argues economic models should take the impact of such narratives into consideration. This paper argues that given investors' limited attention, a few narratives drive security prices at any given point. To test this argument we quantify 347 narratives curated from a large dataset of over 150k digital media sources (roughly 5-7 million articles per week) over the period 2013-2023. We show that stock narrative exposures explain the cross-section of expected stock returns.

The finance literature has studied media coverage mainly as a way to measure investor attention at a particular asset level, be it single-name equity, country equity, currency, or fund. These studies often compare assets with high and low media attention and study the impact on expected return (e.g., Fang and Peress (2009), Solomon (2012), Ozik and Sadka (2013), Froot, Lou, Ozik, Sadka, and Shen (2017), and Calomiris and Mamaysky (2019)), volatility and trading volume (e.g., Tetlock (2007)), or fund flow (e.g., Sirri and Tufano (1998)). The vast majority of studies do not consider the context of the coverage. A few recent exceptions are Engle, Giglio, Kelly, Lee, and Stroebel (2020), Eisdorfer, Lou, Ozik, and Sadka (2023), and Hirshleifer, Mai, and Pukthuanthong (2023) that use textual analysis to gauge overall media coverage to particular themes, climate (the first two aforementioned studies) and war (the latter study). Mai and Pukthuanthong (2021) extract 10 narratives discussed in Shiller (2019) from 7 million New York Times articles over 160 years and demonstrate predictability of market return. Bhargava, Lou, Ozik, Sadka, and Whitmore (2023) analyze 73 themes and introduce a method for estimating asset narrative exposures based on return betas to quantified narratives. Bybee, Kelly, Manela, and Xiu (2023) extract information about the state of the economy from textually analyzing Wall Street Journal articles from 1984 to 2017. Using the same dataset, Bybee, Leland and Kelly, Bryan T. and Su, Yinan (2023) summarizes the news data into risk factors which corresponds to state variables in the Intertemporal CAPM framework.

This paper builds on the approach in Bhargava et al. (2023) to show that narratives are priced in the cross-section of expected stock returns. First, narratives are quantified daily, and a corresponding mimicking-portfolio is built for each narrative as a long-short top-bottom 25 stocks with the highest/lowest corresponding narrative beta (stocks in the Russell 3,000 universe; portfolios are rebalanced monthly; betas are computed via a rolling regression using the most recent 52 weekly observations). Then narrative-mimicking portfolios are shown to exhibit a significant spread in the cross-section of expected return based on the growth in the attention to the narratives. That is, narrative-mimicking portfolios of narratives that have experienced an increase in attention over the recent few months tend to outperform narrative-mimicking portfolios of decreased-attention narratives over subsequent months. This narrative-momentum return suggests investors under-react to narratives that experience an increase in coverage over the intermediate-run.

As a first robustness test, we verify this result is not a manifestation of the well-known price momentum (Jegadeesh and Titman (1993)). That is, we verify that the recent past return of the narrative-mimicking portfolio does not explain the future narrative intensity return spread. Indeed, we construct narrative price-momentum-neutral portfolios, and show they perform as well as without the price momentum control. Recent works demonstrate that many asset-pricing anomalies exhibit significant returns during earnings announcements (e.g., Engelberg, McLean, and Pontiff (2018), Savor and Wilson (2016)). An examination of earnings announcements shows the narrative-momentum returns remain significant while omitting earnings-announcement-month returns.

One unique aspect of the methodology used here is that the exposure of a firm to a given narrative is determined by its return exposure to changes in the coverage intensity of that narrative, and does not rely on the specific media coverage of the firm. This is a strength of the approach because most firms are not covered by the media on any given day. Yet, should a firm be covered by the media, it is plausible that combining firm-specific media coverage and narrative beta may provide further precision of the overall narrative exposure of a firm. To study this, we measure the narrative-specific media coverage of a firm, that is, for each narrative we measure the daily percent of articles mentioning a given firm in the context of the narrative. Indeed, a stronger narrative-momentum return is displayed among firms with above-median coverage per narrative. Notably, a narrative-momentum strategy that uses the level of narrative media coverage instead of narrative beta to form narrative-mimicking portfolios does not produce abnormal returns. Therefore, while the narrative coverage level is informative, narrative beta exposure seems more important.

Perhaps one of the reasons that narrative beta seems important relative to level is that the former includes information about directionality, that is, it uses market prices to determine the firms that stand to benefit from a positive versus a negative discussion of a given narrative (that is beta). A narrative factor is computed as the serial changes in the fraction of negative intensity media coverage of a given narrative across all media coverage at that time. Therefore, a positive (negative) narrative beta of a firm signifies a firm that benefits (loses) from negative news about a narrative. The aggregation of narrative information across the many thousands of sources contributes to the precision of the narrative factor and the resulting firm narrative-beta estimates. The paper provides several examples of narratives, such as COVID-19, inflation, recession, armed conflict, and trade tension, and their corresponding industry group exposures.

Another important consideration is the sourcing of narratives. The concern is that narratives are selected and quantified following a jump in attention, which may induce a forward-looking bias. To alleviate such concerns, we rerun the narrative-mimicking return strategy restricting the set of narratives to a subset of 53 evergreen narratives based on the Journal of Economic Literature (JEL) classification. The results remain robust. We conduct several additional tests. First, we classify narratives into 14 narrative groups (tags), based on common characteristics; for example, inflation and recession both belong to a macroeconomics tag. Each narrative-mimicking Portfolio return is then decomposed into the average return of its tag group and its return in excess of the tag group. While most of the general narrative-momentum return is in excess of the tag groups, tag groups also exhibit a distinct and significant narrative-momentum return. Second, similarly, a sector analysis shows most of the narrative-momentum return is intra-sector, while sectors also display a separate narrative-momentum return. Finally, we perform several tests to establish the robustness of our main results, for example, expanding the universe of narratives to 682 sub-narratives, different portfolio weighting schemes, and subsets of stocks.

The results in this paper suggest that while investors respond to short-run narrative shocks as measured by narrative betas, they under-react to long-run narrative trends, manifesting narrative momentum returns. This is consistent with Samuelson's dictum (Samuelson (1998)), whereby prices are micro-efficient but macro-inefficient; that is, investors spend much time analyzing the impact of new information at the stock level, but they miss aggregate, macroeconomic trends. The implication is that persistent economic narratives can explain asset prices.

The rest of the paper proceeds as follows. Section 2 describes the data. Section 3 presents the main results. Section 4 reports various robustness tests, and Section 5 concludes.

2 Data

The analysis in this paper is based on a dataset of news articles from more than 150k digital media sources over the period 2013–2023. News articles from these sources are classified into media reservoirs: 'General', 'Corporate', 'FX', and 'Country Equity'. Using proprietary natural language processing (NLP) algorithms, the articles are further classified within each media reservoir into 347 narratives, including 53 pre-specified Journal of Economic Literature (JEL) narratives and around 300 additional narratives. The narrative series are provided by MKT MediaStats, LLC.

We use quarterly earnings announcement dates from the Center for Research in Security Prices (CRSP) to determine earnings and non-earnings months, and Compustat for market data. Risk-adjusted returns are computed using the Fama-French factors from Kenneth R. French's Data Library. We use the 11 sectors and 25 industry groups from the MSCI Global Industry Classification Standard (GICS) for each stock. We use yearly analyst forecast revisions from the I/B/E/S Summary History files from 2013 to 2023.

2.1 Narrative Intensities

We consider negative intensities for each narrative. Negative (positive) intensity is the fraction of negative (positive) sentiment articles pertaining to a narrative out of the overall discussion, with a value in [0,1]. Intensities of each narrative are derived from one of the five different media reservoirs: 'General', 'Corporate', 'FX', 'Country Equity', and 'Aggregate'. 'General' uses all articles published in the leading news outlets (e.g., CNN, Fox, NY Times, US Today etc.), 'Corporate' uses articles covering individual companies, 'FX' uses articles covering currencies, 'Country Equity' articles covering country-level indexes, and 'Aggregate' uses the average of the four.

Table 1 shows the date and narrative name that has the highest negative intensity level within each narrative. For example, 'COVID-19' had the highest negative intensity in April 2020 out of all the months the narrative existed. The 'Rank' shows the ranking of the narrative's negative intensity level relative to the ones presented in the table. We require narratives to have intensity levels of at least 0.01 to be included in the table. A total of 177 narratives are shown.

2.2 Correlation across Narrative Reservoirs

In this section, we explore whether certain types of media sources lead others. News covering currencies are more likely to be consumed by sophisticated investors who have the ability and resources to process information quickly. Then currency news might lead the discussion of a narrative against other types of news.

We provide suggestive evidence of a lead-lag relationship across reservoirs by showing how the 'COVID-19' narrative unfolded. Figure 1 shows the time-series intensity levels of the 'COVID-19' narrative from January 2020 to April 2020 for each media reservoir: corporate, country equity, FX, and general. Intensities of corporate and country equity reservoirs developed stronger than general reservoirs at the end of January 2020. Moreover, only after March 2020 did the intensity of general reservoirs surpassed country equity and FX reservoirs. This result suggests that country equity and FX reservoirs might be leading general reservoirs.

Based on the above finding, we document lead-lag correlations of intensity differences for all narratives across different media reservoirs. For each media reservoir, we use the 1-month intensity differences of each of the 347 narratives. Then, we run a pooling regression of the intensity differences on lagged intensity differences. Table 2 (a) and (b) show the contemporaneous and lead-lag correlations of monthly intensity differences across different media reservoirs. The positive cross-correlations in Table 2 (a) show that intensity differences across different media types are significantly correlated contemporaneously. Table 2 (b) shows that 'General' intensity differences has positive correlations with lagged 'Corporate', 'Country Equity', and 'FX' intensity differences in the fourth column. This suggests that narratives from 'Corporate', 'Country Equity', and 'FX' news lead the narratives from 'General' news.

To filter out the autocorrelations of intensity differences, we fit an AR(1) model for each narrative within the media reservoir and recalculate the correlations with the model residuals. Table 2 (c) and (d) show the results. In Table 2 (d), we show that the lead-lag correlation between 'Corporate', 'Country Equity', 'FX', and lagged 'General' intensity differences are strengthened. This implies major news outlets are lagging behind news on individual companies, country indices, and currencies.

For the rest of the paper, we use 'Aggregate' intensities which is the average of the 'Corporate', 'Country Equity', 'General', and 'FX' reservoirs.

2.3 Explaining Market Returns

We explore whether narratives can explain excess market returns. For every 3-year period, we run univariate regressions of market returns on 1-month intensity differences for each of the 347 narratives. Then, we present 5 narratives that have the highest R^2 in each period. Also, adjusted-R^2 are presented from multivariate regressions of market returns on 1-month intensity differences using the 5 selected narratives.

Table 3 shows the regression coefficients and narrative names that were selected from regressing the market return on 1-month intensity differences for each narrative. The market returns are explained by a variety of narratives in each of the 3-year periods, including market uncertainty, market crash, recession, and monetary policy. For example, monetary policy explained the most variation in 2020-2022, implying that the market returns were lower when there was negative news about interest rates.

2.4 Stock-level Narrative Betas

The stock-level narrative betas are estimated as the following:

$$r_{i,w} = \alpha_i + \beta_{n,i}\Delta\text{Int}_{n,w} + \beta_{m,i} r_{m,w} + e_{i,w}$$

where $r_{i,w}$ is the weekly returns for stock i in week w (from this Wednesday to the next Wednesday), $\text{Int}_{n,w}$ is the weekly intensity difference of narrative for stock i in week w, $r_{m,w}$ is the market return

(S&P500 returns from SPY) in week w, and $e_{i,w}$ is the error term. The stock-level beta of stock i is estimated using 52-week rolling window regressions for each week.

We use the last estimated value of $\beta_{n,i}$ from the previous month as stock i's beta exposure to narrative n. For example, if we were currently in March 2020, we used the last estimated beta in February 2020.

Figure 2 provides the time series of example narratives and corresponding average beta deciles of S&P500 stocks by industry group. For example, the left side of Figure 2a shows the 'COVID-19' narrative intensities from January 2020 to April 2023. The right side of Figure 2a shows the top 3 and bottom 3 industry groups with the highest and lowest average beta deciles based on stocks within each industry group. We can see that stocks in Software & Services, Household & Personal Products, and Utilities benefit when negative 'COVID-19' intensities are high. Conversely, stocks in Media & Entertainment, Consumer Services, and Energy are harmed when negative 'COVID-19' intensities are high.

2.5 Narrative-mimicking Portfolios

We construct narrative-mimicking portfolios for each 347 narratives using stocks in the Russell 3000 index. The narrative-mimicking portfolio is a long-short portfolio buying the top 25 stocks with the highest narrative beta and selling the bottom 25 stocks with the lowest narrative beta.

3 Results

3.1 Long-short Narrative Portfolio Returns

To study the cross-section of narrative-mimicking portfolios, we consider three types of intensity differences in our analysis: 1-week (1w), 3-month (3M), and 6-month (6M) intensity differences. The 1-week (3, 6-month) intensity difference is calculated as the average intensity over the past 1 week (the past 3, 6 months) minus the average intensity over the previous 1 week (the previous 3, 6 months).

We construct long-short portfolios from the 347 narrative-mimicking portfolios based on narrative-specific intensity differences. The narrative spread portfolio is constructed by buying the top 10 narrative-mimicking portfolios with the highest increase in intensity differences and selling the bottom 10 narrative-mimicking portfolios with the lowest increase in intensity differences.

Table 4 shows the average returns and alpha against the Fama and French 6-factor model (Fama and French (2018)) for each of the J-month intensity differences and K-month holding horizons. J-month intensity differences are defined as J-month intensity averages minus the previous J-month intensity average. K-month holding horizon is defined as the equal-weighted average of K portfolios, each using 1 to K lags of intensity differences. We find that the narrative spread portfolio has positive returns when sorted by 6-month intensity differences. The 11th row shows that using 6-month intensity differences and 6-month holding horizons yields an alpha of 6.96% per year with a t-statistic of 2.53 in the 2013-2023 sample period. Rows 9 to 12 show that the result is robust to different holding horizons from 1 month to 9 months.

Figure 3 shows the cumulative sums of returns for the narrative spread portfolio using 6-month intensity differences and 6-month holding horizons.

Figure 4 shows the cumulative sums of monthly event time returns 1 to 12 months after portfolio formation for the narrative spread portfolio using 6-month intensity differences and 6-month holding horizons. This figure decomposes the event time returns into the top 10 and bottom 10 narratives. Both sides are almost equally contributing to the event time returns.

This result suggests growth in attention is persistent in the intermediate term. When a particular narrative gains significant attention (when the intensity difference is high), we expect the corresponding narrative-mimicking portfolio to have higher returns in the next period if the attention to the narrative is persistent.

3.2 Momentum Double-sorted Narrative Portfolios

We first test whether price momentum (Jegadeesh and Titman (1993)) drives the narrative spread portfolio returns. We show that the past returns of each narrative-mimicking portfolio do not explain the narrative intensity return spread. Each month, we sort the 347 narrative portfolios into bottom 30%, middle 40%, and top 30% based on their M-month past returns for J-month intensity difference portfolios. Then, we construct a long-short portfolio that buys the top 3 and sells the bottom 3 narratives within each of the 3 momentum sorts. Finally, we report the average returns across the three groups in Table 5. 'M' is the number of past months used to control for momentum, 'J' is the J-month intensity difference, and 'K' is the K-month holding horizon.

The 11th row of Table 5 shows that the 6-month intensity differences and 6-month holding horizon portfolio have significant returns and alphas even after double-sorting by 6-month past returns. Portfolio returns exhibit 5.37% annual returns and 6.38% annual alpha against the Fama and French 6-factor model (Fama and French (2018)) with t-statistics of 2.17 and 2.56 in the 2013-2023 sample, respectively. These results suggest price momentum of narrative-mimicking portfolios is not driving the results.

As an additional robustness check, we present average returns of narrative spread portfolios using the top and bottom 5 and 20 narratives in Table 6, along with the top and bottom 10 baseline result. We confirm that using different numbers of top and bottom narratives does not significantly change the average returns and alphas. The top and bottom 5 portfolio has a higher average return and alpha but lower t-stats than the top and bottom 10 portfolio. The top and bottom 20 portfolio has lower average returns and alpha but similar t-stats with the top and bottom 10 portfolios. In addition, we verify that the top and bottom narratives both contribute to the narrative spread portfolio returns.

The rest of the analysis focuses on the narrative spread portfolio that buys the top 10 and sells the bottom 10 narratives with a 6-month holding horizon based on 6-month intensity differences.

3.3 Subset of Stocks with Different Media Coverage

We test whether a firm's exposure to a given narrative is determined by higher return exposure to the changes in narratives (betas) or by differing levels of media coverage. Media coverage of a firm is defined as the average percentage of articles mentioning a given firm in the context of the narrative in the past 12 months (value in [0,1]). At the end of each month, we divide the stocks into 3 groups: no, low, and high media. No media stocks have media coverage of 0. Low and high media stocks are below and above the median media coverage within the subset of stocks with positive media coverage. Then, we form narrative-mimicking portfolios within each media subset using the top 25 and bottom 25 stocks sorted by narrative-specific betas. Roughly 1900 stocks have no media coverage and 1000 stocks have any media coverage for each month across 347 narratives.

Table 7 shows the results for narrative spread portfolios using the 4 types of narrative-mimicking portfolios. High media coverage stocks exhibit the strongest narrative-momentum return (top-bottom) with 6.98% annual returns and 8.98% annual alphas against the Fama and French 6-factor model (Fama and French (2018)).

In addition, we form narrative-mimicking portfolios based on the top 25 stocks with the highest media coverage using all stocks, regardless of their beta. However, Table 7 shows that the narrative spread portfolio using the highest media coverage stocks as narrative-mimicking portfolios do not exhibit abnormal returns.

3.4 Decomposition into Industry Sectors

We decompose stock-level returns into average sector returns and returns excess of its sector. We classify stocks into 11 sectors: communication services, consumer discretionary, consumer staples, energy, financials, health care, industrials, information technology, materials, real estate, and utilities.

Table 8 shows the decomposition results. Returns excess of the sectors contribute to the return and alpha more than the average sector for the top 10 narratives, whereas sector average returns contribute to the return and alpha more than the excess sector for the bottom 10 narratives. Most of the narrative spread portfolio returns are from intra-sector with an annual return of 4.69% (t-statistic 1.95) and annual alpha of 5.69% (t-statistic 2.31), while sectors also exhibit significant narrative-momentum returns with annual returns of 1.46% (t-statistic 2.59) and annual alpha of 1.51% (t-statistic 2.82).

3.5 Decomposition into Narrative Tags

The narratives are structured as the following: 14 narrative tags, 347 narratives, and 682 sub-narratives. Figure 5 shows examples of each decomposition. The 14 narrative tags can be decomposed into 347 narratives based on common characteristics. For example, inflation and recession both belong to a macroeconomics tag. Then, the 347 narratives are divided into 682 sub-narratives. For example, 'Armed Conflict: Iran' and 'Armed Conflict: Ukraine' are both in the 'Armed Conflict' narrative.

We decompose each narrative-mimicking portfolio return into the average return of its tag group and returns excess of its tag group. Then we check whether the narrative spread portfolio returns are driven by tag or excess tag returns.

Table 9 shows the decomposition results. Returns excess of the tag group contributes to the alpha more than average returns of the tag group for both the top 10 and bottom 10 narratives. However, excess tag and average tag returns both exhibits a narrative-momentum return in the narrative spread portfolio. The excess tag and average tag portfolio have alphas of 5.05% and 1.92% with t-statistics of 2.10 and 2.34, respectively. Figure 6 shows the area plot of cumulative sums of returns decomposed into returns excess of the tag group and average returns of the tag group.

3.6 Subsample Analysis

In this section, we conduct a subsample analysis for the narrative spread portfolio using 6-month intensity differences and 6-month holding horizons in 2013-2017 and 2018-2023. Table 10 shows the average returns and alphas in the first half (2013-2017), second half (2018-2023), and the full sample (2013-2023). Table 10a shows that the narrative spread portfolio returns with equal weights to the top 10 and bottom 10 narratives are stronger in the recent sample from January 2018 to April 2023 with an annual return of 8.10% (t-statistic 2.91) and annual alpha of 9.33% (t-statistic 3.22) during the 64 months. The narrative spread portfolio returns with intensity level weights to the top 10 and bottom 10 narratives increase the returns and alphas across each subsample. The intensity levels are the past J-month average intensity of each narrative. Table 10b shows that the narrative spread portfolio returns and alpha increase to 10.93% (t-statistic 2.99) and 11.76% (t-statistic 3.18) when weighted by intensity levels in the second half of the sample.

4 Robustness Tests

We test several different types of narrative spread portfolios to confirm results are robust. The baseline narrative spread portfolio uses $1 price cutoffs for each stock, 347 narratives, 6-month negative intensity differences as signals, 6-month holding horizons, and equal-weighted returns for the top 10 and bottom 10 narrative-mimicking portfolios. The results are presented in Table 11.

First, we add sub-narratives and use a total of 682 narrative-mimicking portfolios. Sub-narratives are more specific narrative topics. For example, 'Corporate Tax: Tax Increase' is a sub-narrative of the 'Corporate Tax' narrative. The return and alpha are similar at 5.94% and 6.96%, although the t-statistics are lower at 1.68 and 2.01. Removing the $1 price cutoff slightly increases the return and alpha of narrative-momentum returns to 6.46% (t-statistic 2.46) and 7.66% (t-statistic 2.86), respectively. Replacing negative intensity with positive intensity or skipping the most recent month in calculating the 6-month intensity difference signal decreases the narrative spread portfolio return and alpha to 3.92% (t-statistic 1.75) and 4.31% (t-statistic 1.85), but the t-statistic remains significant at the 10% level.

Changing equal weights to intensity level weights for the top and bottom 10 narrative portfolios strengthens the return and alpha of the narrative spread portfolio to 7.90% (t-statistic 2.54) and 9.15% (t-statistic 2.93) per year. Intensity level-weighted '6M' is using the average 6-month past intensity levels for each of the top 10 and bottom 10 narratives. This suggests that investors respond more to narratives that have had more attention recently.

Intensity level-weighted '-median' and '1-' uses the average 6-month past intensity levels for the top 10 but uses (intensity level)$-$(the median of all intensity levels) and 1$-$(intensity level) for the bottom 10 narratives. This tests whether high intensity levels are correlated with higher returns. The corresponding intensity level-weighted narrative spread portfolios have a return of 7.94% (t-statistic 2.54) and 5.97% (t-statistic 2.05) with an alpha of 9.25% (t-statistic 2.93) and 7.51% (t-statistic 2.60), respectively.

We also test whether larger intensity differences lead to higher returns. Intensity difference-weighted narrative spread portfolios use the absolute value of 6-month intensity differences ('6M') as weights and absolute value of (intensity differences)−(the median of all intensity differences) as weights for the top and bottom 10 narrative-mimicking portfolios. The intensity difference-weighted ('6M') narrative spread portfolio has a return and alpha of 6.44% (t-statistic 1.95) and 7.86% (t-statistic 2.36), similar to using median-adjusted intensity differences as weights. This is because the median intensity difference is close to 0.

The narrative spread portfolio using the subset of stocks in the S&P500 has lower returns and alpha estimates of 3.76% and 5.38% but higher t-statistics of 2.63 and 3.85, respectively.

The narrative spread portfolio using 53 evergreen narratives, which are classified by the Journal of Economic Literature (JEL), has similar return and alpha estimates of 5.06% (t-statistic 1.90) and 6.67% (t-statistic 2.48) based on the top and bottom 5 narratives.

We also change equal weights to value weights within the subset of firms that have a valid market cap in our sample. Within the same sample of stocks, we present narrative spread portfolio returns using equal-weighted and value-weighted narrative-mimicking portfolios. The value-weighted narrative spread portfolio has higher returns and alphas of 6.93% and 7.56% with similar t-statistics of 1.94 and 2.06, respectively, compared to the equal-weighted portfolio. This mitigates the concern that small-market-capitalization stocks are the primary drivers of the narrative-momentum returns.

4.1 Persistence of Narratives

In this section, we show that the narratives have a high level of persistence. If narratives are persistent, we can interpret the narrative spread portfolio returns to be driven by under-reaction to narratives. This is because the narrative spread portfolio exploits the momentum in narrative attention. We use impulse responses to measure the persistence of each narrative.

First, for each narrative, we fit an autoregressive model with 12 lags using monthly narrative intensities. Then, we generate monthly responses up to 36 months after a unit shock in time 0. Figure 7 shows generated impulse responses for the 347 narratives including the 53 evergreen narratives. We report the 20% to 80% quantile bounds, median of 347 and 53 narratives, respectively, for each forecast month from 1 to 36. The results show that the top 20% of narratives have an half-life of almost 3 months. Also, the median of 53 evergreen narratives have stronger persistence than the median of 347 narratives. Although the model is estimated in-sample, the figure implies that the persistence in narratives might be driving narrative spread portfolio returns.

4.2 Analyst Revision Double-sorted Narrative Portfolios

We test whether the narrative spread portfolio returns can be explained by analysts revisions in earnings estimates. First, we divide 347 narrative-mimicking portfolios into three levels of past analyst revisions by aggregating the stock-level revisions up to the narrative-mimicking portfolio level. Then, we test whether the spread portfolios perform well using narrative-mimicking portfolios within the three groups.

We define analyst revision as the 6-month rolling average of the monthly change in average analyst forecasts for each stock. We use the 6-month rolling average to account for stocks that do not have changes in analyst forecasts every month. The 6-month analyst revisions are aggregated up to the 347 narrative-mimicking portfolio level. Then, we divide the 347 portfolios into top 30%, middle 40%, and bottom 30% by their average analyst revisions. Finally, we construct the narrative spread portfolio within each analyst revision subset using the top 3 and bottom 3 narrative-mimicking portfolios, sorted by 6-month intensity differences. The three portfolios for the top and bottom account for a total of nine long and short portfolios, respectively, which closely matches the baseline result of ten long and short portfolios.

Table 12 shows the results for narrative spread portfolios within different levels of analyst revisions. Narrative spread portfolios in the high analyst forecast revision subset has an annualized alpha of 10.27% (t-statistic of 2.42), larger than the 4.16% and 3.35% in the low and medium analyst forecast revision subsets. The average narrative spread portfolio returns of each subset has an annual alpha of 5.93%, similar to the baseline result in Table 4.

The results suggest that changes in analyst revisions do not incorporate information from trending or descending narratives, thus analysts under-react to narratives along with investors. If analyst forecast revisions reflect trending and descending narratives, the narrative spread portfolio returns might reflect underreaction to analyst revisions. However, the returns are positive and significant after controlling for analyst revisions. Plus, narrative pricing is most prominent when the analyst forecast revisions are high. Thus, analyst forecast revisions do not seem to take narratives into account.

4.3 Performance of Non-earnings Announcement Months

Recent literature documents that asset pricing anomaly portfolios have significantly higher returns on earnings announcement days compared to non-announcement days (Engelberg et al. (2018)), and firms that are expected to announce earnings have significant abnormal returns (Savor and Wilson (2016)). We test whether earnings announcement returns are driving narrative spread portfolio returns. To construct a non-earnings months portfolio, we drop all stock-months that include an earnings date for each stock and recalculate the narrative spread portfolio with the remaining months.

Table 13 shows narrative spread portfolio returns in non-earnings months and all months. The non-earnings months' returns have slightly higher returns and alpha of 6.37% and 8.18%, with significant t-statistics of 1.90 for the alpha.

4.4 Change in Earnings After Portfolio Formation

We look at earnings changes of stocks selected in the narrative spread portfolio from 1 to 8 quarters (2 years) after portfolio formation to see if trending narratives predict future cash flows. Earnings changes are defined as a firm's quarterly earnings change divided by its total asset in the last quarter. We winsorize earnings changes at the 2% and 98% level to mitigate the effect of extreme values.

First, we select the stocks in the long and short side of the narrative spread portfolio sorted by the 6-month intensity change. Second, we separately take an average the earnings changes of long and short side stocks (25 stocks each). Finally, we take the difference and average across all months to account for the long-short portfolio.

Figure 8 shows the cumulative earnings changes 1 to 8 quarters after portfolio formation, without changing the stock composition. The average spread of cumulative changes in earnings spikes at the third quarter and becomes significantly positive after 6 quarters. The results imply that investors underreact to cash flow information contained in trending and descending narratives.

5 Conclusion

The media can be a useful source of information for understanding security prices because they often directly impact investors' assessment of financial securities. Using media coverage to quantify investor attention to a large set of economic narratives, the results in this paper offer new lens through which asset prices can be studied. Instead of modeling asset returns via exposures to systematic risk factors that often not directly related to an underlying economic factor, one can project asset returns on the space of economic narratives to obtain more direct and interpretable measures of exposures. Such an approach may significantly enhance the understanding of the underlying economic forces driving security prices. As such, there are many practical implications as well, for example, alpha capture, risk modeling, and thematic basket creation.

References

Bhargava, Rajeev, Xiaoxia Lou, Gideon Ozik, Ronnie Sadka, and Travis Whitmore. Quantifying narratives and their impact on financial markets. *The Journal of Portfolio Management*, 49(5): 82–95, 2023.

Bybee, Leland, Bryan T. Kelly, Asaf Manela, and Dacheng Xiu. Business news and business cycles. *Journal of Finance, Forthcoming*, 2023.

Bybee, Leland and Kelly, Bryan T. and Su, Yinan, . Narrative asset pricing: Interpretable systematic risk factors from news text. *The Review of Financial Studies*, 2023.

Calomiris, Charles W and Harry Mamaysky. How news and its context drive risk and returns around the world. *Journal of Financial Economics*, 133(2):299–336, 2019.

Eisdorfer, Assaf, Xiaoxia Lou, Gideon Ozik, and Ronnie Sadka. A market-based measure of environment friendly firms. *Working Paper*, 2023.

Engelberg, Joseph, R David McLean, and Jeffrey Pontiff. Anomalies and news. *The Journal of Finance*, 73(5):1971–2001, 2018.

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel. Hedging climate change news. *The Review of Financial Studies*, 33(3):1184–1216, 2020.

Fama, Eugene F and Kenneth R French. Choosing factors. *Journal of Financial Economics*, 128 (2):234–252, 2018.

Fang, Lily and Joel Peress. Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5):2023–2052, 2009.

Froot, Kenneth, Xiaoxia Lou, Gideon Ozik, Ronnie Sadka, and Siyi Shen. Media reinforcement in international financial markets. 2017.

Hirshleifer, David, Dat Mai, and Kuntara Pukthuanthong. War discourse and the cross section of expected stock returns. Technical report, National Bureau of Economic Research, 2023.

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

Mai, Dat and Kuntara Pukthuanthong. Economic narratives and market outcomes: a semi-supervised topic modeling approach. *Available at SSRN 3990324*, 2021.

Ozik, Gideon and Ronnie Sadka. Media coverage and hedge fund returns. *Financial Analysts Journal*, 69(3):57–75, 2013.

Samuelson, Paul A. Summing up on business cycles: opening address. In *Conference Series-Federal Reserve Bank of Boston*, volume 42, pages 33–36. Federal Reserve Bank of Boston, 1998.

Savor, Pavel and Mungo Wilson. Earnings announcements and systematic risk. *The Journal of Finance*, 71(1):83–138, 2016.

Shiller, Robert J. *Narrative economics: How stories go viral and drive major economic events.* Princeton University Press, 2019.

Sirri, Erik R and Peter Tufano. Costly search and mutual fund flows. *The Journal of Finance*, 53 (5):1589–1622, 1998.

Solomon, David H. Selective publicity and stock prices. *The Journal of Finance*, 67(2):599–638, 2012.

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

Table 1: Narrative Intensity Levels

This table presents the month with the highest negative intensity level ('I') for each narrative. Out of the 347 narratives, we present 177 narratives with intensity levels greater than 0.01. 'Rank' is the ranking of intensity levels out of the 177 months presented in this table. The sample period is from 2012 to 2023.

Date	Name	I	Rank	Date	Name	I	Rank	Date	Name	I	Rank
2012-02	Default Risk	0.023	132	2013-04	Environment	0.078	58	2020-03	Stock Market	0.294	3
2012-03	Aging Insurance	0.016	149	2013-04	Government Money	0.024	130	2020-03	Travel & Leisure	0.020	141
2012-03	Buybacks	0.038	101	2013-07	Housing Market	0.030	121	2020-03	Work From Home	0.035	105
2012-03	Corporate Financing	0.135	28	2013-09	Mortgage	0.051	84	2020-04	COVID-19	0.503	1
2012-04	Corporate Profitability	0.159	23	2013-10	Debt Ceiling	0.059	73	2020-04	Digital Health	0.041	98
2012-04	Natural Gas	0.017	145	2013-11	Typhoons	0.012	168	2020-04	Digital Learning Hardware	0.030	122
2012-05	Asset Management Funds	0.099	43	2013-12	Irrational Exuberance	0.011	171	2020-04	Future Education	0.039	100
2012-05	Cloud Computing	0.032	113	2014-02	Emerging Markets	0.096	47	2020-04	ICU	0.020	139
2012-05	Corporate Governance	0.198	13	2014-09	Digital Payments	0.018	143	2020-04	Public Health	0.415	2
2012-05	Internet Innovation	0.088	51	2014-12	Oil	0.154	24	2020-04	Social Opportunities	0.016	150
2012-05	Millennials	0.059	75	2015-01	Volatility	0.137	27	2020-04	Supply Chain	0.213	9
2012-05	Physical Retail	0.012	164	2015-08	Devaluation	0.045	94	2020-05	Capex	0.026	126
2012-05	Social Media	0.046	90	2015-10	Healthcare Prices	0.012	166	2020-05	ECommerce	0.032	114
2012-05	Valuation	0.058	76	2016-06	International Organizations	0.189	15	2020-05	Online Retail	0.031	116
2012-06	Tsunamis	0.013	160	2016-07	Brexit	0.163	22	2020-06	Civil Unrest	0.112	36
2012-06	Worker Hiring	0.025	127	2016-11	Elections	0.210	11	2020-06	Inequality	0.016	151
2012-07	Dividends	0.010	177	2016-11	Physical Infrastructure Spending	0.030	119	2020-06	Racism	0.049	86
2012-07	Droughts	0.017	144	2016-11	Political Uncertainty	0.084	54	2020-10	Fiscal Stimulus	0.100	42
2012-07	Green Deal	0.045	93	2017-02	Immigration	0.084	56	2020-12	COVID-19 Vaccine	0.108	37
2012-07	Money Market	0.056	78	2017-02	Trump	0.218	5	2021-01	Joe Biden	0.142	25
2012-07	Natural Capital	0.022	137	2017-09	Floods	0.032	112	2021-03	Treasury Bonds	0.087	52
2012-07	Quality Investing	0.011	174	2017-09	Hurricanes	0.088	50	2021-05	Bitcoin	0.049	87
2012-07	Water Stress	0.021	138	2017-09	Natural Disaster	0.104	40	2021-06	Meme Stock	0.016	152
2012-08	Automation	0.019	142	2017-09	Value Investing	0.062	71	2021-10	Build Back Better	0.015	155
2012-08	Smart City Infrastructure	0.011	173	2017-12	Corporate Tax	0.034	107	2021-10	Labor Shortage	0.024	129
2012-09	ECB	0.065	69	2018-03	Gun Laws	0.016	148	2021-10	Shortage	0.040	99
2012-10	Efficient Energy	0.031	118	2018-04	Investor Sentiment	0.068	65	2021-10	Supply Chain Disruption	0.034	106
2012-10	Env Opportunities	0.031	115	2018-05	Ethereum	0.015	156	2021-11	Carbon Emissions	0.011	175
2012-10	Investment Bank	0.027	124	2018-07	Cash Flow	0.052	81	2021-12	Mutation	0.113	35
2012-10	New Energy	0.033	109	2018-07	Gold	0.052	83	2022-02	Buy The Dip	0.012	165
2012-10	Product Safety & Quality	0.014	158	2018-12	US Growth	0.030	120	2022-02	Metaverse	0.013	161
2012-10	Renewable Energy	0.031	117	2019-01	Global Growth	0.045	92	2022-03	Armed Conflict	0.218	6
2013-02	Debasement	0.012	167	2019-03	Momentum Investing	0.027	125	2022-03	Commodities	0.199	12
2013-03	Bankruptcy	0.101	41	2019-05	Meat Prices	0.011	172	2022-03	Fossil Fuels	0.216	8
2013-03	Bonds	0.123	31	2019-07	Monetary Policy	0.172	18	2022-03	Gasoline Price	0.065	68
2013-03	Cannabis	0.023	135	2019-07	Privacy Narrative	0.042	96	2022-03	Wheat	0.015	154
2013-03	Corporate Industry	0.121	34	2019-08	Consumer Spending	0.063	70	2022-06	Crime Rate	0.012	163
2013-03	Crypto	0.069	63	2019-08	International Trade	0.179	17	2022-06	Food Prices	0.050	85
2013-03	Cybersecurity	0.074	59	2019-08	Smart Cities	0.012	162	2022-06	Food Shortage	0.010	176
2013-03	Data Protection	0.015	153	2019-08	Trade Tension	0.210	10	2022-06	Inflation	0.241	4
2013-03	Datafication	0.073	60	2019-08	Yield Curve	0.041	97	2022-06	Stagflation	0.022	136
2013-03	Digital Economy	0.099	45	2019-09	Opportunities in Nutrition & Health	0.011	170	2022-07	Abortions	0.023	134
2013-03	Disruptive Technology	0.108	38	2019-09	Unionizing	0.012	169	2022-07	Business Cycle	0.195	14
2013-03	Earning Season	0.078	57	2019-10	Manufacturing	0.044	95	2022-07	Recession	0.166	21
2013-03	Electric Vehicle	0.070	62	2019-10	Streaming Economy	0.020	140	2022-08	Taiwan Conflict	0.014	159
2013-03	ETFs	0.099	44	2019-12	Aging Disease Treatment	0.014	157	2022-09	Energy Crisis	0.035	104
2013-03	Future Mobility	0.071	61	2019-12	Holiday Shopping	0.024	131	2022-09	Europe Energy Crisis	0.029	123
2013-03	FX Market	0.123	32	2020-01	Drones	0.023	133	2022-09	FED	0.182	16
2013-03	GDP	0.065	67	2020-01	Iran	0.108	39	2022-09	Interest Rate	0.168	19
2013-03	Government Fiscal	0.138	26	2020-01	National Security	0.123	30	2022-12	Personal Consumption	0.047	89
2013-03	Healthcare	0.084	55	2020-01	Strikes	0.059	74	2023-02	Autonomous Technology	0.055	79
2013-03	Personal Finance	0.048	88	2020-03	Aging Society	0.036	103	2023-02	Disinflation	0.017	146
2013-03	Privacy & Data Security	0.061	72	2020-03	Asset Derivatives	0.123	33	2023-02	Earthquakes	0.017	147
2013-03	Retail Investors	0.085	53	2020-03	Education	0.095	48	2023-02	Job Market	0.067	66
2013-03	Taxes	0.167	20	2020-03	Healthcare Infotech	0.033	108	2023-02	Machine Learning	0.025	128
2013-03	Wildfires	0.068	64	2020-03	Liquidity	0.090	49	2023-02	Robotics	0.033	111
2013-04	China Growth	0.052	82	2020-03	Market Crash	0.127	29	2023-02	Robotics & AI	0.036	102
2013-04	Climate Change	0.057	77	2020-03	Market Uncertainty	0.099	46	2023-03	Consumer Financial Protection	0.033	110
2013-04	Climate Change Vulnerability	0.054	80	2020-03	Risk	0.217	7	2023-04	Fintech	0.046	91

Table 2: Correlation of 1-month Narrative Intensity Differences across Media Reservoirs

This table shows the contemporaneous and lagged correlation tables of 1-month intensity differences pooled across different media types. The media types are defined as the following: (1) General - all articles published in the leading news outlets (e.g., CNN, Fox, NY Times, US Today etc.), (2) Corporate - articles covering individual companies, (3) FX - articles covering currencies, (4) Country Equity - articles covering country level indices, and (5) Aggregate - the average of the above 4 media types. Residuals from AR(1) models are used in (c) and (d). 1-month intensity differences are defined as 1-month negative intensity averages minus the previous 1-month negative intensity averages. P-values are reported in the parentheses. The sample period is from 2012 to 2023.

(a) Contemporaneous (Monthly)

	Agg_t	$Corp_t$	$Coun_t$	Gen_t	FX_t
Agg_t	1.00	0.65	0.87	0.72	0.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Corp_t$	0.65	1.00	0.31	0.41	0.29
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Coun_t$	0.87	0.31	1.00	0.52	0.77
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gen_t	0.72	0.41	0.52	1.00	0.52
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
FX_t	0.84	0.29	0.77	0.52	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

(b) Lead-Lag (Monthly)

lag 1	Agg_t	$Corp_t$	$Coun_t$	Gen_t	FX_t
Agg_{t-1}	-0.14	-0.01	-0.24	0.11	-0.19
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Corp_{t-1}$	-0.07	-0.08	-0.09	0.07	-0.05
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Coun_{t-1}$	-0.13	0.02	-0.27	0.15	-0.14
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gen_{t-1}	-0.15	-0.02	-0.20	-0.02	-0.18
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
FX_{t-1}	-0.10	0.04	-0.17	0.14	-0.22
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

(c) Contemporaneous AR(1) (Monthly)

	Agg_t	$Corp_t$	$Coun_t$	Gen_t	FX_t
Agg_t	1.00	0.65	0.86	0.72	0.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Corp_t$	0.65	1.00	0.30	0.40	0.29
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Coun_t$	0.86	0.30	1.00	0.51	0.78
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gen_t	0.72	0.40	0.51	1.00	0.50
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
FX_t	0.84	0.29	0.78	0.50	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

(d) Lead-Lag AR(1) (Monthly)

lag 1	Agg_t	$Corp_t$	$Coun_t$	Gen_t	FX_t
Agg_{t-1}	-0.02	0.05	-0.13	0.17	-0.08
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$Corp_{t-1}$	0.00	0.05	-0.06	0.08	-0.03
	(0.85)	(0.00)	(0.00)	(0.00)	(0.00)
$Coun_{t-1}$	-0.02	0.04	-0.13	0.20	-0.04
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Gen_{t-1}	-0.10	-0.02	-0.16	0.06	-0.13
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
FX_{t-1}	0.01	0.06	-0.06	0.19	-0.08
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Table 3: Univariate Regressions of Market Excess Returns on 1-month Narrative Intensity Differences

This table shows the results for univariate regressions of the market portfolio returns on contemporaneous 1-month intensity differences for 3-year periods in each year starting from 2013. The market return is the CRSP value-weighted return excess of the 1-month T-bill rate from Kenneth R. French's Data Library. The 5 narratives in each univariate regression have the highest adjusted R-square values from monthly univariate regressions of market returns on 1-month intensity differences. The coefficients and t-values are from univariate regressions and the adjusted R-squared are from multivariate regressions of market returns on the 5 narratives. 1-month intensity differences are defined as 1-month negative intensity averages minus the previous 1-month negative intensity averages. T-values are presented in the brackets. The regressions have at least 25 observations for each period. To be included in the sample, narratives are required to be in the top 100 largest average intensity for the corresponding period. The 5 narrative coefficients are presented in the order of the R^2 of univariate regressions. Adjusted R^2 from the multivariate regressions of the market return on the 5 selected narratives are presented. The sample period is from 2013 to 2023.

2013-2015		2014-2016		2015-2017	
Narratives	Coef. [T-value]	Narratives	Coef. [T-value]	Narratives	Coef. [T-value]
Market Uncertainty	-2.77 [-4.64]	Stock Market	-0.83 [-4.38]	Global Growth	-2.83 [-5.70]
FED	-0.92 [-3.87]	China Growth	-2.21 [-4.13]	Stock Market	-0.94 [-5.46]
Interest Rate	-1.07 [-3.65]	Market Crash	-1.10 [-3.62]	China Growth	-2.35 [-4.84]
Stock Market	-0.75 [-3.45]	Global Growth	-1.59 [-3.24]	Market Crash	-1.33 [-4.83]
Risk	-1.10 [-3.41]	Volatility	-0.80 [-3.04]	Business Cycle	-1.58 [-4.76]
Adj. R^2	0.42	Adj. R^2	0.32	Adj. R^2	0.53
2016-2018		2017-2019		2018-2020	
Narratives	Coef. [T-value]	Narratives	Coef. [T-value]	Narratives	Coef. [T-value]
Recession	-2.78 [-3.46]	Market Crash	-2.70 [-4.88]	Market Crash	-2.74 [-7.29]
Business Cycle	-1.84 [-3.40]	Stock Market	-1.10 [-3.56]	Stock Market	-1.81 [-5.63]
Market Crash	-1.66 [-3.37]	Market Uncertainty	-2.34 [-3.35]	Risk	-2.37 [-5.62]
US Growth	-3.65 [-3.12]	Risk	-1.25 [-3.30]	Education	-4.60 [-5.11]
Global Growth	-3.05 [-3.12]	Investor Sentiment	-1.89 [-3.04]	Volatility	-2.00 [-4.80]
Adj. R^2	0.33	Adj. R^2	0.43	Adj. R^2	0.59
2019-2021		2020-2022		2021-2023	
Narratives	Coef. [T-value]	Narratives	Coef. [T-value]	Narratives	Coef. [T-value]
Market Crash	-2.47 [-6.93]	Monetary Policy	-1.96 [-4.70]	Monetary Policy	-2.19 [-3.38]
Stock Market	-1.88 [-6.19]	Market Crash	-2.25 [-4.19]	Natural Disaster	-3.46 [-2.52]
Volatility	-2.19 [-5.79]	Education	-4.66 [-4.19]	Floods	-9.74 [-2.47]
Risk	-2.21 [-5.66]	Stock Market	-1.92 [-3.93]	FED	-1.32 [-2.46]
Education	-4.40 [-5.44]	Asset Derivatives	-4.33 [-3.86]	Interest Rate	-1.60 [-2.44]
Adj. R^2	0.56	Adj. R^2	0.36	Adj. R^2	0.34

Table 4: Top and Bottom 10 Narrative-mimicking Portfolio Returns sorted by J-month Narrative Intensity Differences and K-month Holding Horizons

This table shows portfolio returns and alpha for different J-month narrative intensity differences and K-month holding horizons. J-month intensity differences are defined as J-month negative intensity averages minus the previous J-month negative intensity averages. K-month holding horizon is defined as the equal-weighted average of K portfolios, each using 1 to K lags of intensity differences. All portfolios buy the top 10 and sell the bottom 10 narratives based on J-month intensity differences. All portfolios skip the most recent day when calculating intensity differences. Yearly returns and alphas against the Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. Monthly t-values are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from April 2013 to April 2023 (121 months), September 2013 to April 2023 (116 months), March 2014 to April 2023 (110 months), and September 2014 to April 2023 (104 months), respectively, depending on J-month intensity differences availability.

	J	K	Return (%)	T-value	Alpha (%)	T-value	N
1	1W	1M	-0.31	-0.10	-1.32	-0.44	121
2	1W	3M	0.55	0.20	-0.70	-0.26	121
3	1W	6M	0.22	0.08	-1.04	-0.38	121
4	1W	9M	-0.22	-0.08	-1.27	-0.47	121
5	3M	1M	1.54	0.58	0.03	0.01	116
6	3M	3M	-0.14	-0.06	-0.68	-0.29	116
7	3M	6M	1.32	0.72	1.19	0.65	116
8	3M	9M	2.70	1.91	3.06	2.10	116
9	6M	1M	6.61	2.03	6.89	2.02	110
10	6M	3M	6.64	2.19	7.69	2.46	110
11	6M	6M	5.77	2.13	6.96	2.53	110
12	6M	9M	3.97	1.75	4.41	1.95	110
13	9M	1M	6.42	1.83	8.22	2.28	104
14	9M	3M	5.37	1.60	6.64	1.93	104
15	9M	6M	3.55	1.21	4.07	1.37	104
16	9M	9M	2.46	0.96	2.66	1.03	104

Table 5: Top and Bottom 3 Narrative-mimicking Portfolio Returns sorted by J-month Intensity Differences using K-month Holding Horizons, within M-month Past Return Sorts

This table shows portfolio returns and alpha for different J-month narrative intensity differences and K-month holding horizons, double-sorted by momentum portfolios. First, portfolios are sorted into top 30%, middle 40%, and bottom 30% portfolios based on M-month past returns. Then, each of the three portfolios buys the top 3 and sells the bottom 3 narratives based on J-month intensity differences. Finally, the average returns across the three groups are reported. J-month intensity differences are defined as J-month intensity averages minus the previous J-month intensity averages. K-month holding horizon is defined as the equal-weighted average of K portfolios, each using 1 to K lags of intensity differences. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from April 2013 to April 2023 (121 months), September 2013 to April 2023 (116 months), March 2014 to April 2023 (110 months), and September 2014 to April 2023 (104 months), respectively, depending on J-month intensity differences availability.

	M	J	K	Return (%)	T-value	Alpha (%)	T-value	N
1	1M	1W	1M	-1.43	-0.49	-2.64	-0.92	121
2	1M	1W	3M	-0.07	-0.03	-1.62	-0.63	121
3	1M	1W	6M	-0.36	-0.13	-2.10	-0.79	121
4	1M	1W	9M	-0.06	-0.02	-1.66	-0.65	121
5	3M	3M	1M	2.71	0.95	1.22	0.43	116
6	3M	3M	3M	-0.52	-0.23	-1.03	-0.45	116
7	3M	3M	6M	0.53	0.29	0.20	0.11	116
8	3M	3M	9M	2.25	1.59	2.36	1.62	116
9	6M	6M	1M	3.92	1.23	4.42	1.35	110
10	6M	6M	3M	5.44	1.95	6.29	2.23	110
11	6M	6M	6M	5.37	2.17	6.38	2.56	110
12	6M	6M	9M	3.99	1.90	4.32	2.06	110
13	9M	9M	1M	6.77	2.01	7.81	2.25	104
14	9M	9M	3M	6.56	2.06	7.69	2.36	104
15	9M	9M	6M	4.75	1.72	5.45	1.94	104
16	9M	9M	9M	3.52	1.42	4.06	1.62	104

Table 6: Different Numbers of Top and Bottom Narrative-mimicking Portfolio Returns sorted by 6-month Intensity Differences

This table shows the equal-weighted returns of narrative-mimicking portfolios sorted by 6-month narrative intensity differences. Portfolio returns of top and bottom 5, 10, and 20 narrative-mimicking portfolios are reported. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. The top minus bottom returns are divided by 2 to maintain a long-short portfolio gross leverage of 2. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from March 2014 to April 2023 (110 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Top - Bottom 5 narratives				
Top	6.45	1.10	9.27	1.55
Middle	0.86	0.64	0.70	0.50
Bottom	-5.90	-1.05	-5.58	-0.99
Top - Bottom (200% Gross)	6.18	1.76	7.42	2.09
Top - Bottom 10 narratives				
Top	6.05	1.25	8.70	1.75
Middle	0.88	0.69	0.67	0.50
Bottom	-5.50	-1.05	-5.23	-0.98
Top - Bottom (200% Gross)	5.77	2.13	6.96	2.53
Top - Bottom 20 narratives				
Top	4.93	1.27	6.86	1.74
Middle	0.80	0.66	0.56	0.44
Bottom	-2.63	-0.63	-2.84	-0.66
Top - Bottom (200% Gross)	3.78	1.88	4.85	2.35

Table 7: Top and Bottom 10 Narrative-mimicking Portfolio Returns sorted by 6-month Intensity Differences using Stocks with Different Media Coverage

This table shows the equal-weighted returns of narrative-mimicking portfolios with no media coverage, low media coverage, high media coverage stocks sorted by 6-month intensity, and using long-only portfolios with high media coverage stocks. Stocks are split into no, low, and high media coverage sorts by the level of media exposure to the corresponding narrative in the last 12 months. Stocks that have 0 coverage is classified as no media, below and above median coverage (excluding the no-media stocks) are classified as low and high media. The long-only narrative-mimicking portfolios use the top 25 media coverage stocks for each narrative. Long- or short-only portfolio returns are reported excess of the value-weighted Russell 3000 portfolio. The top minus bottom returns are divided by 2 to maintain a long-short portfolio gross leverage of 2. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2017 to 2023, and portfolio returns are calculated from February 2018 to February 2023 (61 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
No Media Coverage Stocks				
Top - Bottom 10 narratives				
Top	7.12	2.11	8.42	2.27
Bottom	-0.86	-0.23	0.47	0.12
Top - Bottom (200% Gross)	3.99	1.67	3.98	1.53
Low Media Coverage Stocks				
Top - Bottom 10 narratives				
Top	3.72	0.98	6.46	1.63
Bottom	-3.11	-0.63	-0.00	-0.00
Top - Bottom (200% Gross)	3.41	1.39	3.23	1.34
High Media Coverage Stocks				
Top - Bottom 10 narratives				
Top	3.99	0.84	8.78	1.97
Bottom	-9.98	-2.43	-9.18	-2.23
Top - Bottom (200% Gross)	6.98	2.81	8.98	3.50
Top 25 Media Coverage Stocks				
Top - Bottom 10 narratives				
Top	-0.95	-0.20	-0.87	-0.36
Bottom	-2.78	-0.52	-0.57	-0.18
Top - Bottom	1.82	0.53	-0.30	-0.08

Table 8: Decomposition of Narrative-mimicking Portfolio Returns into Excess and Average Sector Returns

This table shows the decomposition of the top 10 minus bottom 10 narrative-mimicking portfolio returns based on 6-month narrative intensity differences into excess and average sector returns. Stocks are required to have valid sector information to be included in the sample. The 11 sectors include communication services, consumer discretionary, consumer staples, energy, financials, health care, industrials, information technology, materials, real estate, and utilities. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from March 2014 to April 2023 (110 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Top 10 narratives				
Excess Sector	5.13	1.12	7.56	1.61
Average Sector	0.67	0.59	0.64	0.59
Total	5.80	1.17	8.20	1.63
Bottom 10 narratives				
Excess Sector	-4.26	-0.93	-3.82	-0.82
Average Sector	-2.25	-2.07	-2.39	-2.06
Total	-6.51	-1.33	-6.20	-1.25
Top - Bottom 10 narratives				
Excess Sector	4.69	1.95	5.69	2.31
Average Sector	1.46	2.59	1.51	2.82
Total	6.16	2.31	7.20	2.67

Table 9: Decomposition of Narrative-mimicking Portfolio Returns into Excess and Average Narrative Tag Returns

This table shows the decomposition of the top 10 minus bottom 10 narrative-mimicking portfolio returns based on 6-month narrative intensity differences into excess and average narrative tag returns. Narrative tag is a broader definition that classifies narratives. For example, Cloud Computing, 3D Printing, and Robotics & AI are in the 'Tech' narrative tag. Narrative exposure returns are decomposed into average and excess narrative tag returns. Average narrative tag returns are the average narrative exposure return within each tag. Excess tag returns are narrative-mimicking portfolio returns minus average narrative-mimicking portfolio returns within each narrative tag. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from March 2014 to April 2023 (110 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Top 10 narratives				
Excess Tag	4.76	1.45	6.69	1.98
Average Tag	1.29	0.56	2.00	0.83
Total	6.05	1.25	8.70	1.75
Bottom 10 narratives				
Excess Tag	-2.74	-0.77	-3.40	-0.96
Average Tag	-2.76	-1.13	-1.83	-0.74
Total	-5.50	-1.05	-5.23	-0.98
Top - Bottom 10 narratives				
Excess Tag	3.75	1.57	5.05	2.10
Average Tag	2.02	2.59	1.92	2.34
Total	5.77	2.13	6.96	2.53

Table 10: Subsample Analysis for Top and Bottom 10 Narrative-mimicking Portfolio Returns using J-month Intensity Differences and K-month Holding Horizons

This table shows equal-weighted and intensity level-weighted portfolio returns and alpha for 1 week, 3-month, and 6-month intensity differences and 1, 3, and 6-month holding horizons, respectively, within the 2013-2023, 2013-2017, and 2018-2023 period. Panel (a) shows returns using equal-weighted top 10 (bottom 10) narrative portfolios and Panel (b) shows returns using past J-month intensity level-weighted top 10 (bottom 10) narrative portfolios. J-month intensity differences are defined as J-month negative intensity averages minus the previous J-month negative intensity averages. All portfolios buy the top 10 and sell the bottom 10 narratives based on J-month intensity differences. A portfolio with K-month holding horizon is defined as the equal-weighted average of K portfolios, each using 1 to K lags of intensity differences. All portfolios skip the most recent day when calculating intensity differences. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. Monthly t-values are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from April 2013, September 2013, March 2014, respectively to December 2017 for the first-half depending on J-month intensity differences availability, and January 2018 to April 2023 for the second-half.

(a) Equal-weighted Top and Bottom 10

Period	J	K	Return (%)	T-value	Alpha (%)	T-value	N
First-half [2013-2017]	1w	1M	-0.21	-0.05	1.64	0.32	57
	3M	3M	-4.56	-1.21	-5.78	-1.39	52
	6M	6M	2.53	0.49	1.00	0.20	46
Second-half [2018-2023]	1w	1M	-0.40	-0.10	-2.08	-0.55	64
	3M	3M	3.45	1.27	2.66	1.07	64
	6M	6M	8.10	2.91	9.33	3.22	64
Full sample [2013-2023]	1w	1M	-0.31	-0.10	-1.32	-0.44	121
	3M	3M	-0.14	-0.06	-0.68	-0.29	116
	6M	6M	5.77	2.13	6.96	2.53	110

(b) Intensity-weighted Top and Bottom 10

Period	J	K	Return (%)	T-value	Alpha (%)	T-value	N
First-half [2013-2017]	1w	1M	2.39	0.38	5.87	0.85	57
	3M	3M	-3.06	-0.75	-4.55	-1.02	52
	6M	6M	3.68	0.71	1.79	0.35	46
Second-half [2018-2023]	1w	1M	-3.53	-0.64	-4.86	-0.82	64
	3M	3M	3.88	1.30	4.27	1.51	64
	6M	6M	10.93	2.99	11.76	3.18	64
Full sample [2013-2023]	1w	1M	-0.74	-0.18	-1.13	-0.26	121
	3M	3M	0.77	0.31	0.78	0.31	116
	6M	6M	7.90	2.59	9.15	2.97	110

Table 11: Robustness Tests for the Top and Bottom 10 Narrative-mimicking Portfolio Returns using 6-month Intensity Differences and 6-month Holding Horizons

This table shows portfolio returns and alpha for 6-month intensity differences and 6-month holding horizons using different types of specifications. Portfolio construction follows Table 4. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. Monthly t-values are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023. '682 Sub-narratives' uses sub-narratives (i.e., 'Corporate Tax: Tax Increase' for narrative 'Corporate Tax') in addition to the 347 narratives. 'Positive intensity' uses positive intensity differences instead of negative intensity differences as signals. 'Skipping 1 Month' skips the most recent month and uses the past 5 month minus the previous 6 month in estimating the signal. 'No Price Cutoff' does not use the $1 stock price cutoff in the previous month. 'Intensity level-weighted' uses intensity levels using past 6-month average in weighting the top 10 but uses either the same past 6-month average, 1 minus the past 6-month average, or absolute value of past 6-month average relative to its median for the bottom 10 narrative-mimicking portfolios. 'Intensity difference-weighted' uses 6-month intensity differences in weighting the top 10 but uses either the same absolute value of 6-month intensity difference or absolute value of 6-month intensity difference relative to its median for the bottom 10 narrative-mimicking portfolios. 'S&P 500 Subset' uses stocks that were in the S&P500 in the previous month. The sample period is from 2013 to 2023, and portfolio returns are calculated from March 2014 to April 2023 (110 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Baseline	5.77	2.13	6.96	2.53
682 Sub-narratives	5.94	1.68	7.07	2.01
Positive Intensity	3.94	1.73	4.24	1.80
Skipping 1 Month	3.92	1.75	4.31	1.85
No Price Cutoff	6.46	2.46	7.66	2.86
Intensity level-weighted (6M)	7.90	2.59	9.15	2.97
Intensity level-weighted (-median)	7.94	2.54	9.25	2.93
Intensity level-weighted (1-)	5.97	2.05	7.51	2.60
Intensity difference-weighted (6M)	6.44	1.95	7.86	2.36
Intensity difference-weighted (-median)	6.45	1.95	7.87	2.36
S&P500 Subset	3.76	2.63	5.38	3.85
Evergreen Narrative Subset (Top 10)	3.55	1.77	4.51	2.20
Evergreen Narrative Subset (Top 5)	4.28	1.67	5.53	2.13
Market Cap Subset EW	4.50	1.80	5.58	2.19
Market Cap Subset VW	6.93	1.94	7.56	2.06

Table 12: Top and Bottom 3 Narrative-mimicking Portfolio Returns sorted by J-month Intensity Differences using K-month Holding Horizons, within 6-month Past Analyst Revision Sorts

This table shows portfolio returns and alpha for 6-month narrative intensity differences and 6-month holding horizons, double-sorted by past analyst forecast revisions. First, portfolios are sorted into top 30%, middle 40%, and bottom 30% portfolios based on average 6-month rolling analyst forecast revisions of each individual stock in the narrative portfolio scaled by market capitalization. Then, portfolios buys the top 3 and sells the bottom 3 narratives based on 6-month intensity differences within each of the three analyst forecast groups. Finally, average returns for each of the three groups are reported. J-month intensity differences are defined as J-month intensity averages minus the previous J-month intensity averages. K-month holding horizon is defined as the equal-weighted average of K portfolios, each using 1 to K lags of intensity differences. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. The sample period is from 2013 to 2023.

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Low Analyst Forecast Revisions				
Top - Bottom 3 narratives				
Top - Bottom (200% Gross)	4.06	1.04	4.16	1.01
Medium Analyst Forecast Revisions				
Top - Bottom 3 narratives				
Top - Bottom (200% Gross)	1.14	0.34	3.35	1.01
High Analyst Forecast Revisions				
Top - Bottom 3 narratives				
Top - Bottom (200% Gross)	9.54	2.35	10.27	2.42
Average of All Analyst Forecast Revisions				
Top - Bottom 9 narratives				
Top - Bottom (200% Gross)	4.91	1.86	5.93	2.20

Table 13: Decomposition of Narrative Spread Portfolio Returns into Earnings and Non-earnings Announcement Month Returns

This table shows the results of the top 10 minus bottom 10 narrative-mimicking portfolio returns based on 6-month intensity differences using stocks on non-earnings months. Non-earnings month returns are recalculated after dropping all stock-months that includes an earnings date for each stock. Yearly returns and alphas against Fama and French 6-factor model (Fama and French (2018)) are reported in percentages. T-values of monthly returns are reported. 'N' is the number of months that are used in a portfolio. The sample period is from 2013 to 2023, and portfolio returns are calculated from March 2014 to April 2023 (110 months).

Portfolio	Return (%)	T-value	Alpha (%)	T-value
Top 10 narratives				
Non-earnings months	8.31	1.30	10.81	1.67
All months	6.05	1.25	8.70	1.75
Bottom 10 narratives				
Non-earnings months	-5.73	-0.89	-5.37	-0.83
All months	-5.50	-1.05	-5.23	-0.98
Top - Bottom 10 narratives				
Non-earnings months	6.37	1.50	8.18	1.90
All months	5.77	2.13	6.96	2.53

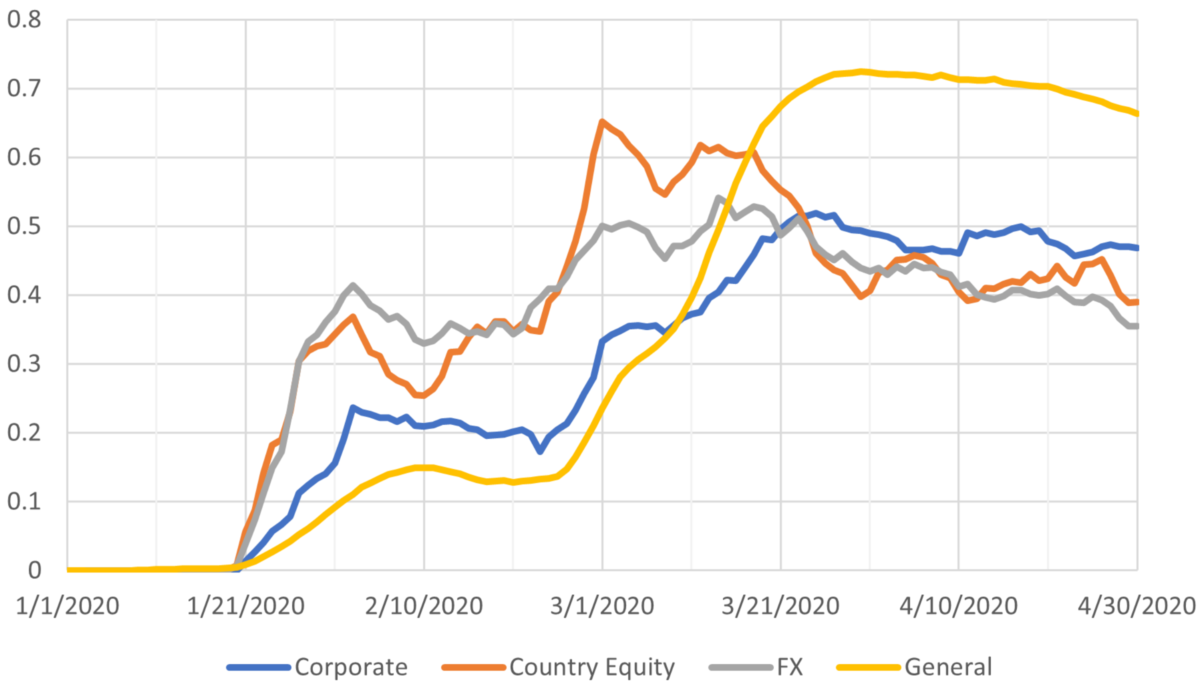
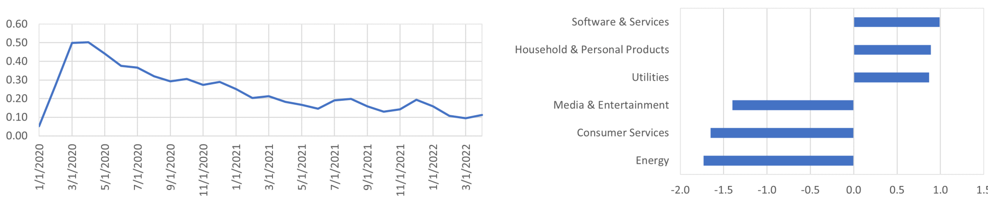
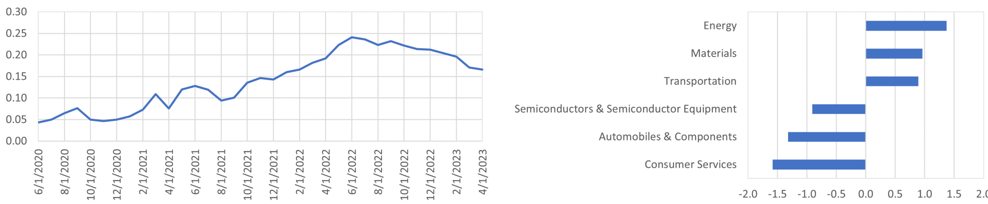


Figure 1: Time-series Negative Intensity of 'COVID-19' Narrative for each Media Reservoir

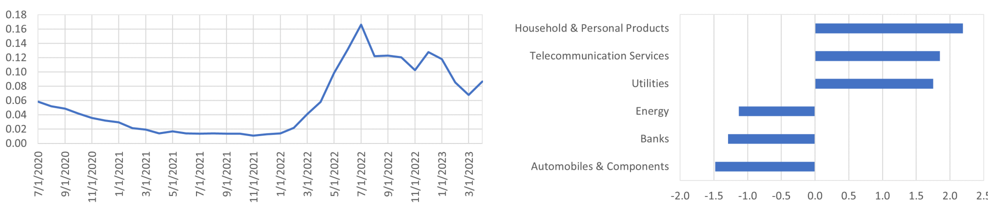
This figure shows the daily time-series negative intensity levels of the 'COVID-19' narrative for each media reservoir: corporate, country equity, FX, and general. The sample period is from January 2020 to April 2020.



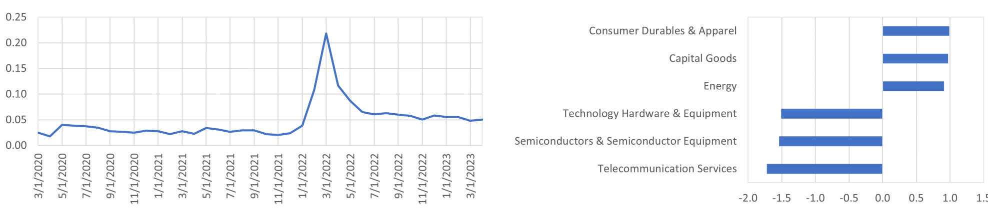
(a) COVID-19



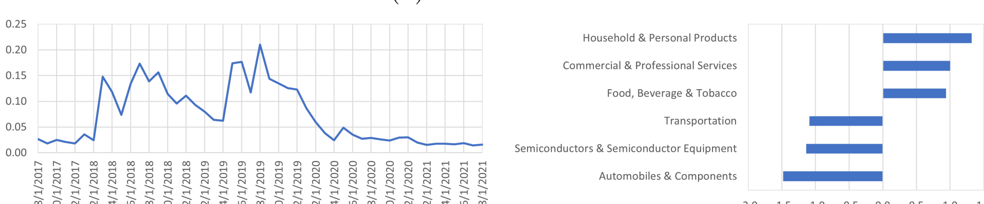
(b) Inflation



(c) Recession



(d) Armed Conflict



(e) Trade Tension

Figure 2: Average Beta Deciles for each Narrative in each Industry Group with S&P500 Stocks 1 (before/after 2y of highest intensity month average)

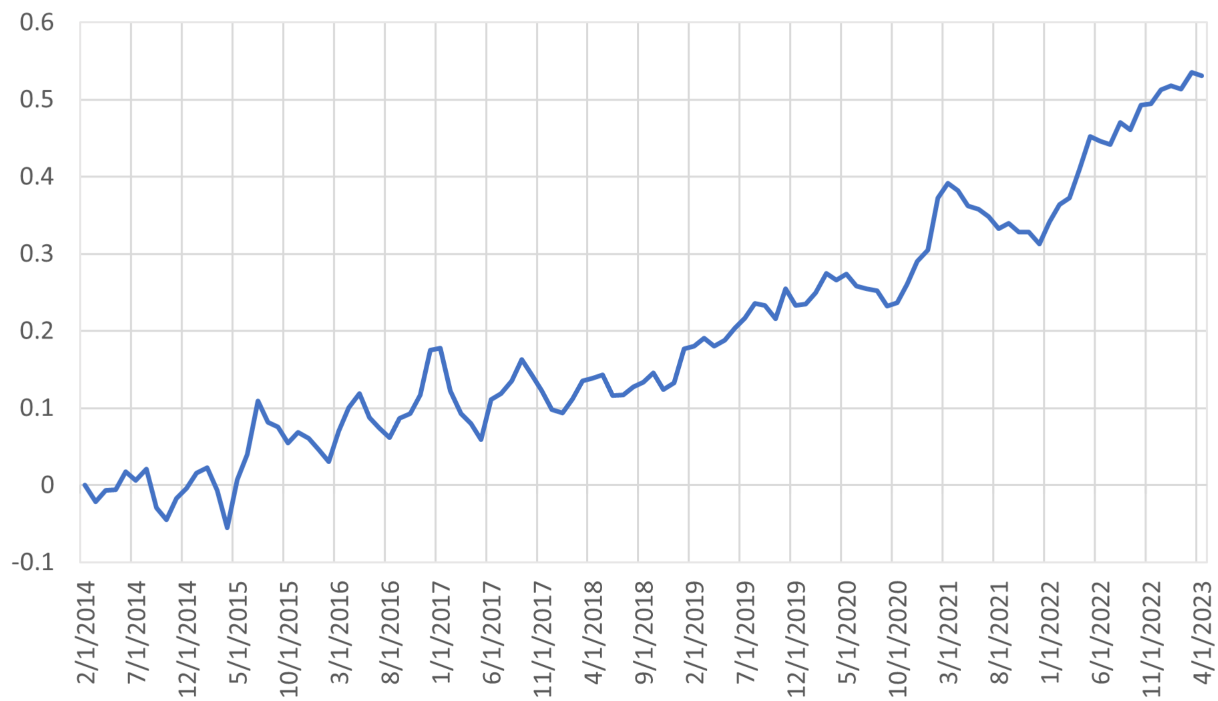


Figure 3: Time-series Cumulative Returns of the Narrative Portfolio (6M-6M)

This figure shows the time-series cumulative sums of monthly returns of the portfolios using 6-month intensity differences and 6-month holding horizons. Portfolio construction follows Table 4. The sample period is 2013-2023.

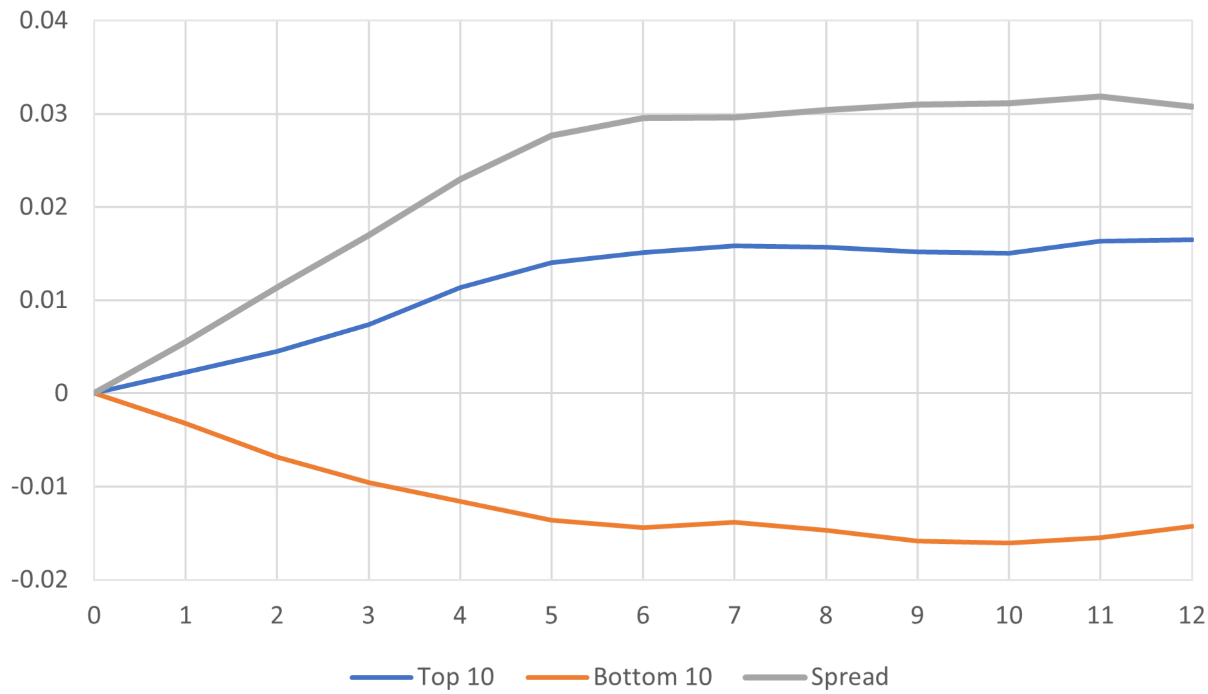


Figure 4: Event Time Cumulative Monthly Returns after the 6M-6M Portfolio Formation

This figure shows the event time cumulative sums of monthly returns for the 6M-6M portfolio. 'Top 10' and 'Bottom 10' are the top and bottom 10 narrative portfolios sorted by 6-month intensity differences. 'Spread' is the long-short portfolio. The sample period is 2013-2023.

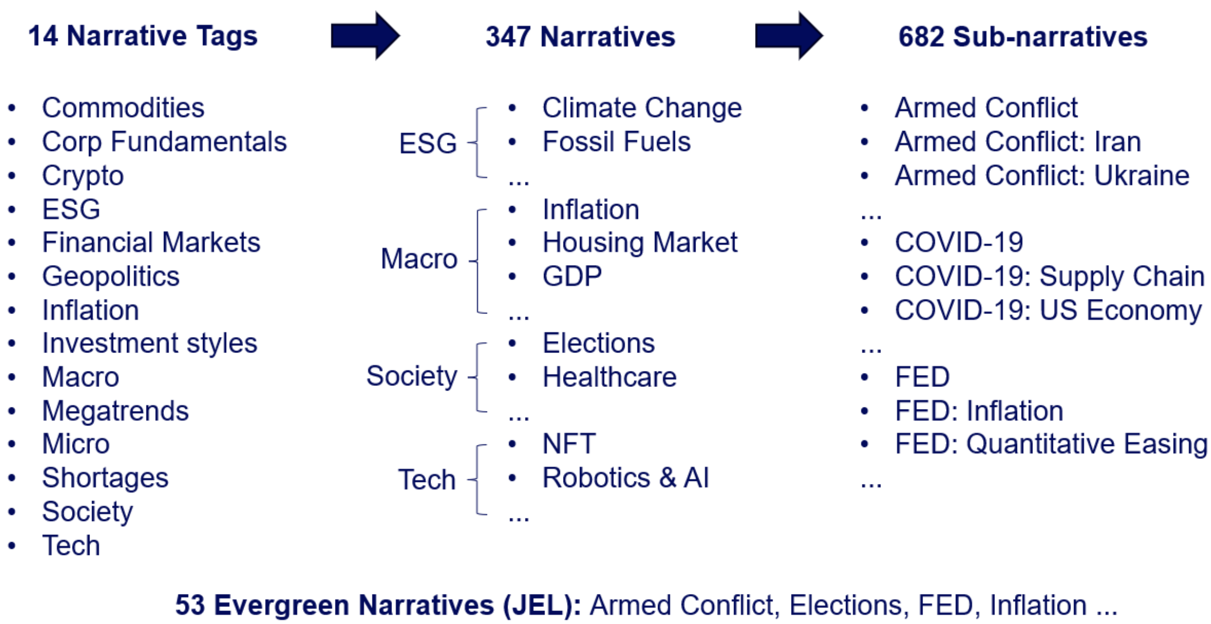


Figure 5: Decomposition of Narrative Tags into Narratives and Sub-narratives

This figure shows examples of narrative tags decomposed into narratives and sub-narratives. The 14 narrative tags can be divided into 347 narratives which include the 53 evergreen narratives, where the 347 narratives can be divided again into 682 sub-narratives. The 53 evergreen narratives correspond to the Journal of Economic Literature (JEL) narratives.

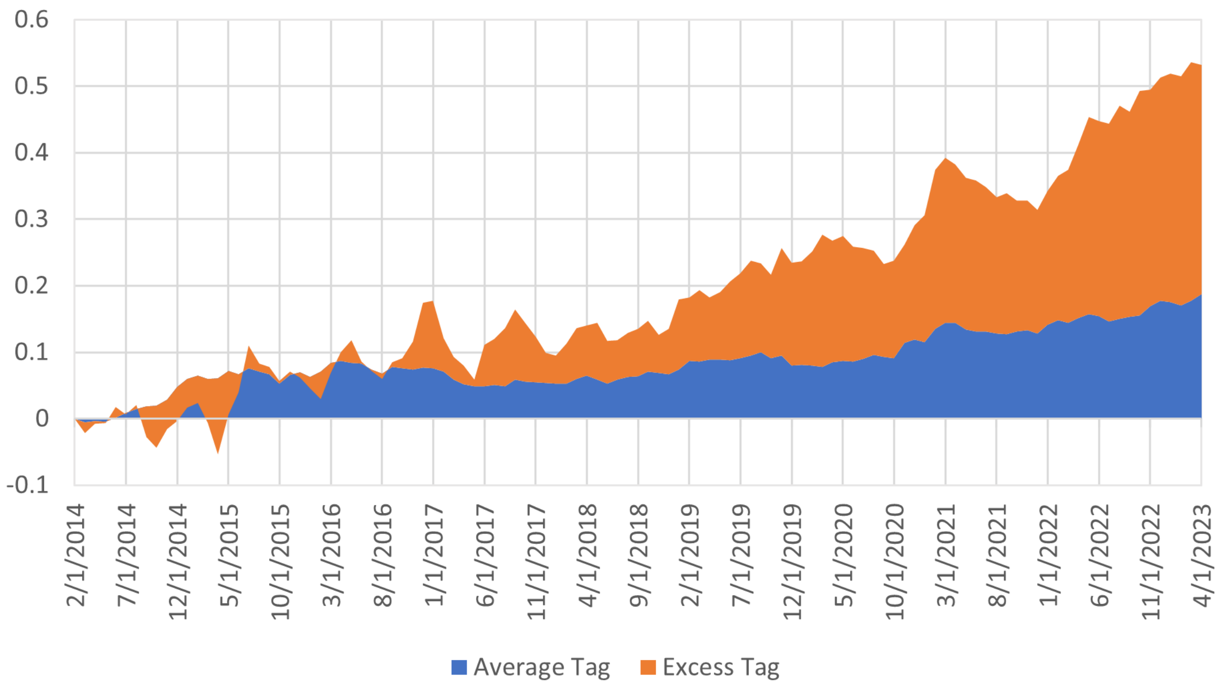


Figure 6: Decomposition of the 6M-6M Portfolio into Average and Excess Tag Returns

This figure shows the time-series cumulative sums of monthly 6M-6M portfolio returns decomposed into average and excess tag returns. Narrative tag portfolio construction follows Table 7. The sample period is 2013-2023.

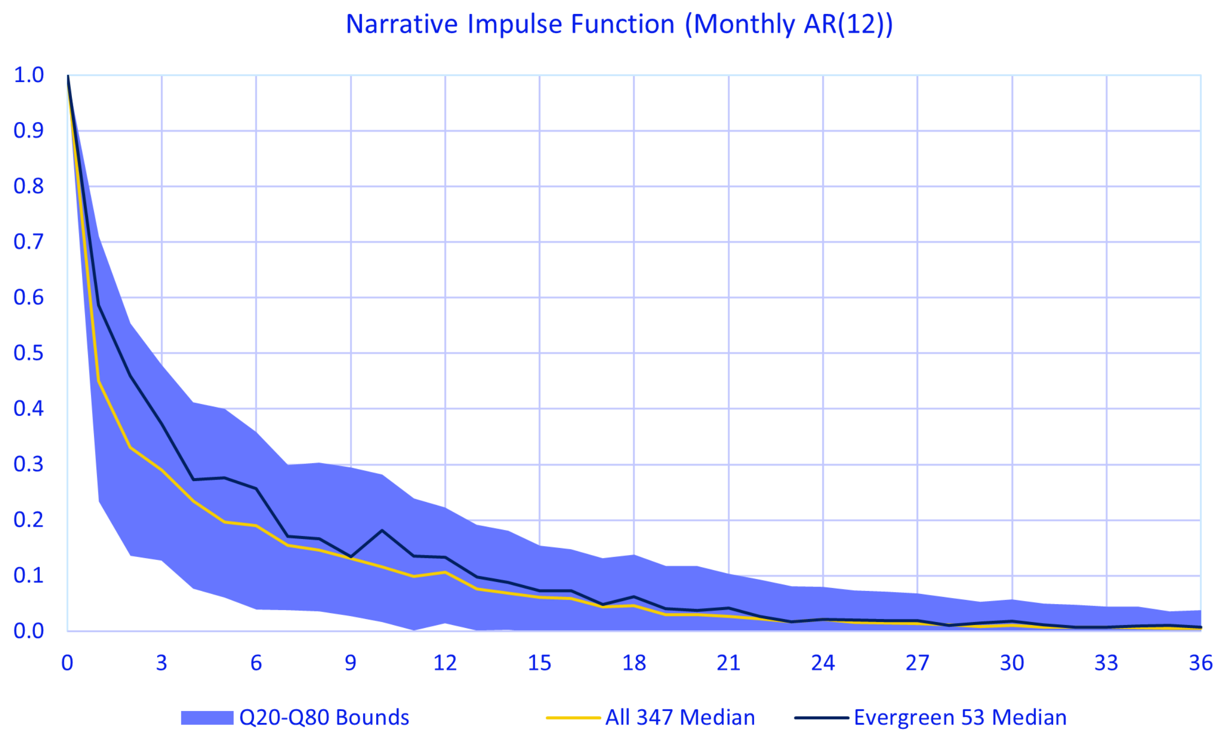


Figure 7: Impulse Responses for the 347 Narratives Including the 53 Evergreen Narratives

This figure shows the impulse responses of 347 narratives including the 53 evergreen narratives using a in-sample monthly AR(12) model and responses up to 36 months. The Q20-Q80 bound represents the impulse responses between 20% and 80% quantiles among the 347 narratives for each forecast month. The yellow (black) line represents the median impulse response among the the 346 narratives (53 evergreen narratives) for each forecast month. The sample period is 2013-2023.

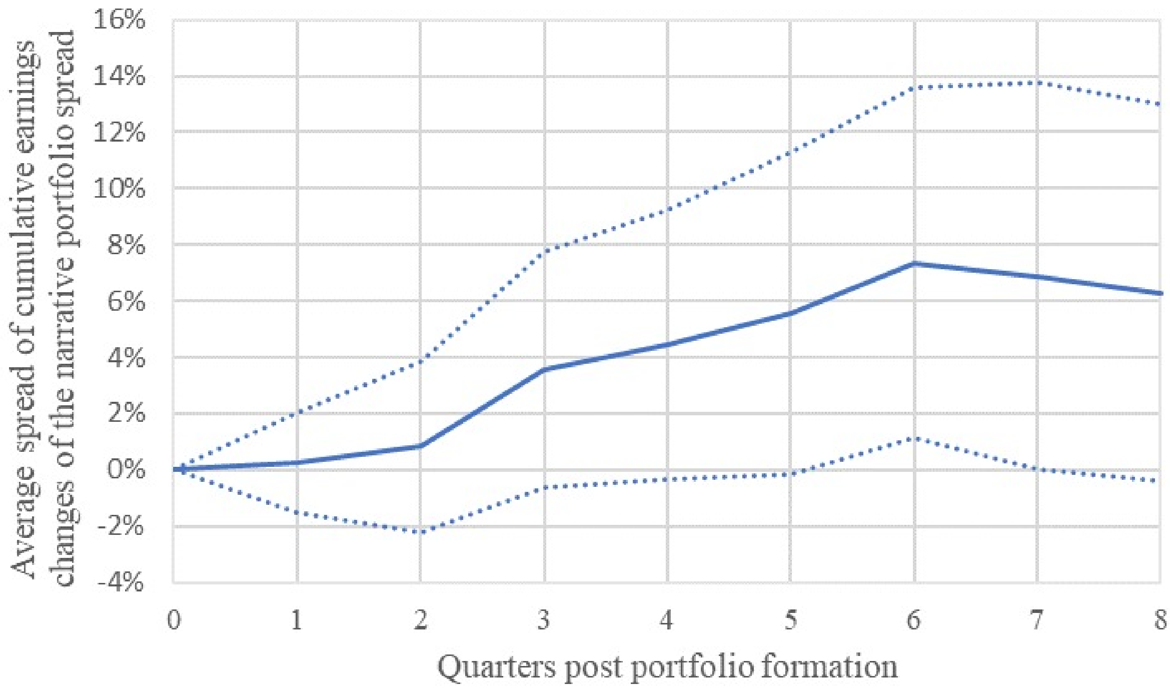


Figure 8: Average Spread of Cumulative Quarterly Earnings Changes Post Portfolio Formation

This figure shows the time-series cumulative sums of quarterly earnings changes of stocks in the top 10 minus bottom 10 narrative portfolios selected by past 6-month intensity differences. The sample period is 2013-2023.