Characterizing Fake News Targeting Corporations

Ke Zhou[1,2], Sanja Šćepanović[1], Daniele Quercia[1,3]

[1] Nokia Bell Labs, UK
[2] University of Nottingham, UK
[3] King's College London, UK
ke.zhou@nokia-bell-labs.com

Abstract

Misinformation proliferates in the online sphere, with evident impacts on the political and social realms, influencing democratic discourse and posing risks to public health and safety. The corporate world is also a prime target for fake news dissemination. While recent studies have attempted to characterize corporate misinformation and its effects on companies, their findings often suffer from limitations due to qualitative or narrative approaches and a narrow focus on specific industries. To address this gap, we conducted an analysis utilizing social media quantitative methods and crowd-sourcing studies to investigate corporate misinformation across a diverse array of industries within the S&P 500 companies. Our study reveals that corporate misinformation encompasses topics such as products, politics, and societal issues. We discovered companies affected by fake news also get reputable news coverage but less social media attention, leading to heightened negativity in social media comments, diminished stock growth, and increased stress mentions among employee reviews. Additionally, we observe that a company is not targeted by fake news all the time, but there are particular times when a critical mass of fake news emerges. These findings hold significant implications for regulators, business leaders, and investors, emphasizing the necessity to vigilantly monitor the escalating phenomenon of corporate misinformation.

1 Introduction

Emerging research in economics and political science has raised concerns about the erosion of democracy by misinformation on social media (Allcott and Gentzkow 2017). Similarly, concerns about misinformation effects on public health are increasing (Naeem, Bhatti, and Khan 2021; Melchior and Oliveira 2022). Yet another sphere experiencing threat due to misinformation is corporations and capital markets (Adriani 2019; Fox 2020; Jahng 2021). There are various types of corporate fake news depending on whether the news content has the intention to deceive and harm (Park et al. 2020). When fake news has a low intention to harm and deceive, the news contents might just provide inaccurate information due to honest mistakes or negligence (e.g., inaccurate product release leakage). However, some fake news has the intention to harm the company but not to deceive, i.e., promoting factual but negative news used to inflict harm on the competing organization (e.g., layoffs), while others focus on both deceiving and harming, i.e., manipulation of information (e.g., company executive scandal) that purposefully aims to mislead and misinform the audience/customers. All those types of corporate fake news may have an impact to different extents on a company's reputation and its capital markets. Research on corporate fake news is limited due to a lack of data, incorporating narrative (Adriani 2019; Camacho and Lozano 2020), qualitative (Serazio 2021), and survey-based approaches (Cheng and Chen 2020; Bronnenberg, Dubé, and Sanders 2020), along with a focus on small parts of the corporate world (Kogan, Moskowitz, and Niessner 2019; Chung, Zhang, and Pan 2022).

This study aims to characterize and provide empirical evidence of the determinants of corporate fake news across S&P 500 companies, using a large-scale quantitative analysis on social media data and crowd-sourced validation. The study addresses three research questions (RQs):

(RQ1) How should corporate fake news articles be categorized?

(RQ2) Which companies are targeted by fake news?

(RQ3) When are companies targeted by fake news?

To answer those research questions, our study makes four main contributions:

- By combining social media posts from Reddit, fake news links, company stock performance, and company reviews from Glassdoor, we *created a unique dataset* for investigating corporate misinformation. We *developed ten metrics* to capture the characteristics of companies and the misinformation they were subject to (§3).
- We *conducted a user study to fully categorize the news articles* mentioning companies in the context of misinformation. We found that a two-level taxonomy best describes these misinformation articles and its first level includes products, politics, corporate affairs, societal issues, and brand. Fake news articles mentioning high-growth companies tend to be about products and societal issues, while those mentioning limited-growth companies tend to be about politics (§4, *RQ1*).
- We *characterized the relationship between misinformation and companies*. We found that companies heavily targeted by fake news often receive coverage from reputable news publishers but garner less attention on social

media. However, this not only leads to more negative sentiments in public comments on social media but also links to decreased stock market growth, as well as increased stress indicators among employee reviews within these targeted companies.

- We found a company is not targeted by fake news all the time – there are particular times in which a critical mass of fake news about the company emerges (§6, *RQ3*).

2 Related Work

Misinformation, evident historically like propaganda in WWII (Fox 2020), thrives now, termed a 'perfect storm' (Starbird 2021; Fox 2020; Edge 2021). Influenced by: (i) easy content production via online platforms; (ii) social media's virality, especially for fake news (Vosoughi, Roy, and Aral 2018); and (iii) advancements in deep learning, enabling deepfakes (Westerlund 2019). Research shows its impact on democratic and political discourse (Pickard 2019; Garimella et al. 2018), public health (Swire-Thompson and Lazer 2019; Verma et al. 2022), journalism (Waisbord 2018), and social divisions (Cover, Haw, and Thompson 2022). Misinformation also threatens the corporate world (Jahng 2021; Adriani 2019). For instance, false ties between COVID-19 and 5G led to destroying 5G towers (Ahmed et al. 2020; Moshood, Shittu, and Abidin 2020).

Fake news significantly impacts companies, primarily through brand reputation (Jahng 2021; Di Domenico and Visentin 2020; Mills and Robson 2019; Castellani and Berton 2017). Nearly all professionals view damage to brand reputation as a risk, with 80% facing at least one crisis due to fake news (Camacho and Lozano 2020). Financial implications include market manipulation, sales decline, and share value loss (Lin 2016; Adriani 2019; Castellani and Berton 2017). Major corporations like Apple or Amazon could suffer billion-dollar losses from successful misinformation campaigns (Adriani 2019). It has been shown in a small-scale study (Xu 2021) that fake news perpetrators can come both from people external to the company but also from employees themselves if they want to excerpt pressure on the company. Prompt response to corporate misinformation can mitigate damage to reputation and stock price (Mills and Robson 2019; Bronnenberg, Dubé, and Sanders 2020).

While initial evidence about the effects of fake news on companies starts to mount (PWC 2022), current research uses mostly narrative (Adriani 2019; Camacho and Lozano 2020), qualitative (Serazio 2021), and survey-based approaches (Cheng and Chen 2020), or focuses on a small part of the corporate world (e.g., financial market (Kogan, Moskowitz, and Niessner 2019)). Quantitative large-scale research is still nascent owing to the methodological challenges in quantifying fake news (Al-Rawi and Fakida 2021) and their effects on companies (Adriani 2019; Jahng 2021; Xu 2021; Di Domenico and Visentin 2020). Our current research addresses this research gap by offering insights into the fake news targeting S&P 500 companies shared on Reddit from 2016 to 2019. Despite recent research findings indicating limited influence of misinformation on political outcomes (Eady et al. 2023), it's vital to recognize and understand the unique dynamics in the corporate context. Misinformation exposure here could significantly affect internal or external corporate reputation and, consequently, stock valuation. Unlike entrenched political beliefs, perceptions of a company's offerings may be more susceptible to false information among consumers, employees, and investors.

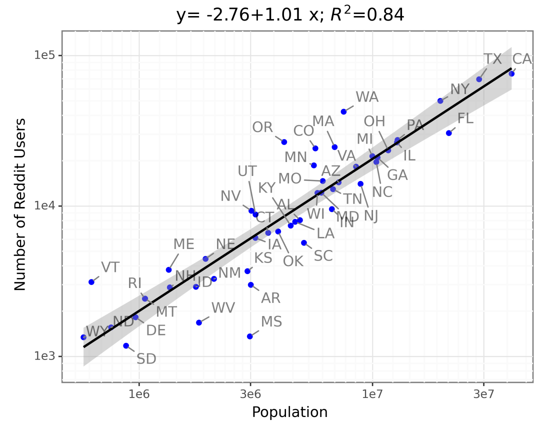

Figure 1: Reddit users versus U.S. state population.

3 Data and Methods

To examine and compare the susceptibility of S&P 500 companies to misinformation, we collated four types of data. We collected Reddit comments mentioning those companies along with news articles (§3.1). The credibility of these articles was evaluated utilizing established fact-checking sources (§3.2). It is possible that a company's likelihood of being targeted by fake news may simply be related to its level of newsworthiness. This news coverage may stem from an external interest in the company, such as financial performance as measured by stock growth (§3.3), or internal factors within the company. Internal company affairs have so far been harder to characterize. To do so, we analyzed data from a reputable company review website, Glassdoor (§3.4). Finally, using these four data sources, we computed ten metrics for each company (§3.5).

3.1 Reddit Comments with Company News

We chose Reddit to collect public perception of news articles related to companies. This is because: i) Reddit serves as a social news aggregation site, ideal for studying news distribution and public response, and ii) unlike typical social networks (e.g., Twitter), Reddit lacks friends or followers, allowing for a more indicative reflection of mainstream news production and consumption, free from social network biases like echo chambers or filter bubbles. We used publicly available Pushshift API (Baumgartner et al. 2020) to download the Reddit comments from January 2016 to December 2019. We included all comments from all public and quarantined subreddits (e.g., r/the_Donald). Since we studied S&P 500, we chose to focus only on the users who are from the U.S.

Reddit does not explicitly provide user location. However, it is possible to obtain reliable estimates (Balsamo, Bajardi, and Panisson 2019). Starting from a list of 2,844 location-based subreddits that map to one of the U.S. cities or states (e.g., r/newyork, r/california), we listed the users with at least 5 comments in those subreddits. If they had posted in multiple states, we assigned the user to the state with the majority of comments. In this way, we placed ∼3M users in one of the U.S. states. The number of Reddit users per state scaled linearly with the state's population (β = 1.01, R^2=0.84) showing a good geographic representativeness (Figure 1).

Between 2016 and 2019, these users authored 1.4 billion comments. Of these comments, 8.2M (0.6%) contained news links, of which 1.4M (16%) mentioned also an S&P 500 company (Table 1). Those comments contained 11.3M/1.3M total/unique links and had 262 (52.4%) of the S&P 500 companies mentioned (Wikipedia 2022). Finally, we classified a Reddit comment as a fake news comment if it contained a URL to a domain that appeared in the list of fake news domains, which was defined using the credibility sources described next (§3.2). In total, there were 12.8K fake Reddit comments mentioning S&P 500 companies. We can observe the median number of Reddit comments per month discussing corporate fake news is 408 (Figure 2).

3.2 News Credibility Sources

We established the credibility of news at the level of the publisher (or domain). Specifically, all articles published on a domain deemed as fake or of low credibility were also considered fake or not reputable. To determine domain credibility, we employed a commonly referenced list of news domains in misinformation research (Bozarth, Saraf, and Budak 2020). Specifically, we labeled domains as reputable using lists compiled by Alexa (Alexa.com 2022), Media Bias/Fact Check (MBSFC 2022), and Vargo et al.(Vargo, Guo, and Amazeen 2018). This resulted in a total of $8,900$ reputable news sites. Non-reputable domains were labeled using Media Bias/Fact Check and four additional sources: Zimdars list (Zimdars 2016), PolitiFact (Gillin 2018), the Daily Dot (Dot 2022), and Allcott, et al (Allcott, Gentzkow, and Yu 2019). By merging the labels from each of these five sources, we classified a domain as fake if it had a history of publishing fabricated news articles. This resulted in a total of 933 fake news domains. Upon this classification, we found that 1.4 million Reddit comments included mentions of $145.5K/39.7K$ total/unique fake news articles, as well as mentions of $114 (22.8\%)$ of the S&P 500 companies (as outlined in Table 1).

3.3 Company Stock Market Data

We collated the S&P 500 stock market data, including the monthly adjusted closing prices from 2016 until 2019 for each company, from the Yahoo Finance portal (Yahoo Finance portal 2022). The stock data quantifies external perception of the company from an investor perspective.

3.4 Glassdoor Company Reviews

To understand internal employees' perception of their company, we acquired data regarding internal corporate affairs

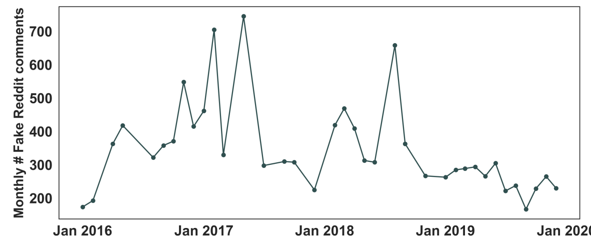

Figure 2: Monthly # of Reddit comments containing both a fake link and a company name (median: 408, standard deviation: 136). The peaks observed in early 2017 and late 2018 were mainly caused by news about Google and Facebook. They were accused of not complying with the law that requests the removal of harmful content (Facebook) and using politically biased algorithms (Google).

from 440K reviews of the S&P 500 companies posted during the twelve years 2008-2020 on Glassdoor, a popular company reviewing site (Table 2). In 2021, there were $50M$ monthly visitors on the platform, and $70M$ reviews about $1.3M$ companies. Current and former employees of companies describe their job experience, ranging from interviews to salaries to workplace culture. All the reviews we collected were from U.S.-based employees. We also obtained the employee ratings of each company.

3.5 Computing Company Metrics

For each company, we collated ten metrics, with formulas listed in Table 3, and described below.

`#Reddit comments`: we used the total number of Reddit comments discussing a company as a proxy for its popularity among the general public. To make sure that this measure does not capture simply the size of the company, we also tested a corrected metric for popularity described next.

`#Reddit comments per capita`: we calculated the total number of Reddit comments divided by the total number of Glassdoor reviews, which serves as a reasonable proxy for company's employee count.

`#reputable news`: we calculated the total number of news articles mentioning a company in reputable news outlets, which generally captures a company's newsworthiness.

`external reputation score`: we computed `Q-score` (Freeman and Dart 1993) (one minus the fraction of the number of comments with a negative sentiment to the total number of comments mentioning the company) as a proxy for public perception of a company. To measure the sentiment of a comment, we used the Flair library (Akbik et al. 2019), and applied it to *all* the comments mentioning the company, and not only those with a news link.

`Glassdoor rating`: we calculated the average rating from the company reviews on Glassdoor as a proxy for the internal perception, i.e., the company's popularity among the (former) employees.

`internal employee stress score`: we calculated the percentage of Glassdoor reviews discussing stress

Type of comments	# comments	# users	# total links	# unique links	# domains	# companies
comments by geolocated users in the U.S.	1.4B	3M				
company comments	11.1M	795K				262
company comments with news	1.4M	322K	11.3M	1.3M	5.1K	208
company comments with fake news	12.8K	6.3K	145.5K	39.7K	316	114

Table 1: Statistics about Reddit comments mentioning companies and news links present in the comments.

# total reviews	# US-based reviews	# companies	avg (min-max) rating	avg (min-max) stess score	time period
713K	440K	399	3.37 (1.62-5)	1.11% (0-9.52%)	2008-2020

Table 2: Statistics about Glassdoor reviews of S&P 500 companies.

Metric	Formula	Description
`# Reddit comments (C, T)`	#Reddit comments mentioning C during period T	public popularity
`# Reddit comments per capita (C, T)`	$\frac{\text{\#Reddit comments mentioning C during period T}}{\text{\#Glassdoor reviews of C during period T}}$	corrected public popularity
`# reputable news (C, T)`	#reputable news articles mentioning C during period T	newsworthiness
`external reputation score (C, T)`	$1 - \frac{\text{\#negative Reddit comments mentioning C in period T}}{\text{\#total Reddit comments mentioning C in period T}}$	positive/neutral mentions on social media
`stock valuation (C, T)`	adjusted closing prices at end of period T for company C	external investor perception
`stock growth (C)`	$\frac{\text{stock valuation (C, 2019)}}{\text{stock valuation (C, 2014)}}$	external investor perception
`Glassdoor rating (C, T)`	Glassdoor overall rating of company C within period T	internal perception
`internal employee stress score (C, T)`	$\frac{\text{\#Glassdoor stress reviews for company C in period T}}{\text{\#total Glassdoor reviews for company C in period T}}$	stress mentions by employees
`pFake (C, T)`	$\frac{\text{\#fake Reddit comments mentioning C during period T}}{\text{\#Reddit comments mentioning C during period T}}$	probability of `C` being targeted by fake news
`misinformation shock (C, T)`	# (pFake(C, T_month)>pFake(C, T)+2std(pFake(C, T)))	number of concentrated misinformation attacks

Table 3: List of company metrics.

at work as a proxy for stress mentions by employees. To extract work stress mentions from reviews, we used the MedDL entity extraction (Scepanovic et al. 2020), which uses contextual embeddings and a deep BiLSTM-CRF sequence labeling architecture (we used the default parameters as specified in (Scepanovic et al. 2020)). The model was pre-trained and evaluated on a labeled dataset of 2K Reddit posts that have in total 4.5K medical mentions called MedRed (Scepanovic et al. 2020). The state-of-the-art MedDL method achieved a $F1$-score of .85 when extracting symptoms (including stress mentions).

`stock growth`: we calculated the ratio of the `stock valuation` (average adjusted closing prices) in 2019 and 2014 as a proxy for external investor perception of the company. We chose to focus on the growth of the companies between these two years for two reasons: (i) it encompasses the posting dates of the Reddit comments in our data, thus providing a relevant time frame for examination. Additionally, by selecting a stock growth starting date of 2014, two years prior to the earliest Reddit comments in our dataset (from 2016), it allows for the determination of the company's growth rate as it would have been known to Reddit users at that time; (ii) by focusing on long-term growth rather than short-term growth, we increase the robustness of the metric to potential influences of external events such as market manipulation or incidental growth/decline (for 2021).

`pFake`: we computed the probability of news articles about the company being fake as the ratio of fake to total Reddit comments with articles discussing the company.

`misinformation shock`: we determined all the months with `pFake`(C, T_{month}) two standard deviations above its mean within the entire period T for that company C. Prior studies have shown that misinformation can be spread in a coordinated manner, where several outlets cover the same fake news in a short period of time (Ng and Taeihagh 2021; Muzykant et al. 2021). To identify such occurrences, we identified periods of 'misinformation shocks' for companies, i.e., these concentrated periods (months) when companies are targeted by a significant number of fake news articles. These 'misinformation shocks' correspond to relatively large coordinated misinformation events.

We calculated each of these metrics for the entire time period T_{all} between 2016 and 2019. The five metrics that had heavy-tailed distributions (`# Reddit comments`, `# Reddit comments per capita`, `# reputable news articles`, `stock growth` and `pFake`) were log-transformed in our analysis.

4 How to Categorize Fake News Articles?

Before understanding the determinants of corporate fake news, we would like to first characterize them, specifically addressing (*RQ1*): *How should corporate fake news articles be categorized?* Previous studies have highlighted the importance of identifying the intentions of news outlets in the dissemination of fake news as a crucial step in combating its proliferation (Tandoc Jr 2019; Kalsnes 2018).

To address this question, we employed a multi-faceted approach. First, we drew upon existing literature to develop a taxonomy of fake news article categories. This was subsequently expanded through the use of manually-labeled annotations on a subset of articles, thereby identifying additional categories (§4.1). Second, we conducted a crowd-sourced study in which all articles were annotated with the identified categories during periods of misinformation shocks (§4.2), and reported the resulting categories in §4.3.

4.1 Identifying Categories

Previous studies have explored the intentions and motivations behind the creation and dissemination of fake news. For instance, Zhou and Zafarani (2020) categorized fake news intentions as 'mislead', 'entertain', and 'undefined', while others like (Kalsnes 2018) classified motivations as 'political', 'financial', and 'social'. Despite these endeavors, it remains unclear how these general intentions can be translated into a more nuanced and fine-grained taxonomy specifically for corporate fake news articles.

To address this gap in the literature, we conducted a qualitative study using a subset of randomly sampled 100 corporate fake news articles. Three annotators were instructed to independently annotate the categories of these articles using the intention taxonomies from prior literature (Pennycook and Rand 2021; Lazer et al. 2018; Kalsnes 2018; Zhou and Zafarani 2020). The annotators then discussed their labels to reach a final agreement on an extended taxonomy.

As shown in Table 4 (Level 1), we found that articles were primarily categorized as covering products, politics, corporate affairs, societal issues, and brands. Additionally, we added a category of "Others" to capture any additional categories that may arise. Within each high-level level-1 category (e.g., product), the news articles were further categorized into more fine-grained level-2 categories (e.g., product issues relating to quality), as shown in Table 4 (Level 2). In total, there are 19 level-2 categories.

4.2 Annotating Articles with Categories

We then used the manually-labelled taxonomy to label a larger subset of N=1500 salient news articles with a crowd-sourcing approach.

Identifying articles during misinformation shocks. In order to identify articles during periods of misinformation, a representative sample of fake news articles was selected. Specifically, articles that covered the companies of interest and were shared during "misinformation shocks" (as defined in §3) were considered representative. To mitigate the potential for the most popular companies to dominate the results, a random selection process was employed as follows:

Level 1	Level 2	% Articles in Category	Morality
Product	**Issues - Quality**	9.9%	Purity (-)
	Launches	8.4%	Loyalty (-)
	Issues - Supply	2.7%	Purity (-)
	Issues - Security	2.2%	Authority (-)
Politics	**Lobbying**	9.7%	Loyalty (-)
	Boycotting	6.7%	Harm (+)
	Regulation	4.4%	Fairness (-)
Corporate Affairs	Strategy	8.0%	Fairness (-)
	Lay off	4.0%	Harm (+)
	Acquisition	2.7%	Authority (-)
	Work conditions	2.2%	Fairness (-)
	Corruption	0.9%	Harm (+)
Societal Issues	Conspiracy	7.5%	Harm (+)
	Minorities - Racial Issues	6.5%	Fairness (-)
	Censorship	1.8%	Fairness (-)
	Minorities - LGBT	0.9%	Fairness (-)
Brand	Criticism	6.4%	Purity (-)
	Promotion	5.8%	Loyalty (+)
Others	**Others**	9.7%	Fairness (-)

Table 4: Companies that are targeted by fake news tend to be mentioned in articles mainly covering products, politics, corporate affairs, societal issues and brands. The top 3 and bottom 3 categories that are widely covered by news outlets are in bold and italic respectively. For each category, the top associated moral dimension (with the highest absolute morality score) is shown in the "Morality" column.

(a) we randomly selected a company C; (b) we extracted a random set of $k = 50$ articles about company C shared during the months of misinformation shocks; (c) we iterated the previous two steps until N news articles were collected for annotations. To facilitate comparison, a random sample of trustworthy news articles from the same peak periods was also obtained. This enabled the examination of both fake and trustworthy news articles.

Crowd-sourcing data collection. To annotate the news articles according to the established taxonomy, a crowd-sourcing approach was employed utilizing Amazon Mechanical Turk. The data collection process consisted of the following three main steps. First, participants completed a short pre-study questionnaire in which they self-reported their age and gender. Second, participants were instructed to read the news article in detail. Third, participants were instructed to respond to three questions:

Q1: Which company does this article mainly cover?
This question served to collect the company names and also served as a trap question to assess annotation quality. Participants who failed more than one trap question had their annotations removed from the analysis.

Q2: Under which section of corporate coverage would you classify this article?
This question collected the categories of the representative articles in Table 4. If the assessors could not find any applicable category, they could add their own categorization through a free-text form under the option "Others".

Q3: How would you judge the company based on the article (assuming it is true)? Please pick five words each from a different moral dimension.

Harm	Fairness	Loyalty	Authority	Purity
harmful	unjust	disloyal	disobedient	indecent
violent	discriminatory	traitor	defiant	obscene
caring	*fair*	*devoted*	*lawful*	*decent*
protective	*impartial*	*loyal*	*respectful*	*virtuous*

Table 5: Moral dimensions along which our participants judged a company mentioned in news articles.

This question assessed how the company would be judged in the article, were it to be true, along five moral dimensions: harm, fairness, loyalty, authority, and purity (Graham et al. 2011). The participants were asked to pick five words each from a different moral dimension. Table 5 illustrates the positive (in *italic*) and negative words used for those moral dimensions for the participants to pick from (Graham et al. 2011). We then calculated the morality score of each moral dimension by counting the corresponding number of positive words selected minus the number of negative words selected within that dimension.

A total of 103 participants were involved in this crowdsourcing study, with each new article being evaluated by three participants. On average, each participant evaluated around 44 news articles. The final categorization of each article was determined using a majority vote approach.

4.3 Popular News During Misinformation Shocks

Categories of fake news articles. The top three high-level (Level 1) categories for corporate misinformation are: *Product*, *Politics* and *Societal* issues. Among lower-level (Level 2), the top categories are *Product–Issues–Quality*, followed by *Politics–Lobbying* and *Product–Launches* (see Table 4 for a complete breakdown). Notably, the "Others" category accounted for only 9.7% of fake news articles, indicating that over 90% of fake news articles were covered by the manually established taxonomy. This partially validates the effectiveness of the taxonomy in categorizing corporate fake news articles.

In Table 4, under the 'Morality" column, we present the top moral dimension with the highest absolute morality score for each fake news category. For instance, in news articles that cover *Product-Issues-Quality,* people tend to perceive the discussed companies as morally wrong (i.e., negative *purity*, frequently judging them as "indecent" and "obscene".). Overall, by tracking morality scores for all fake news articles, we found that companies that are targeted by fake news tend to be judged as being unfair (i.e., *unjust* and *discriminatory* in `fairness`), morally wrong (i.e., *indecent* and *obscene* in `purity`), and harmful (i.e., *harmful* and *violent* in `harm`). We also compared the morality scores across different news categories. It is not surprising that, for news articles that cover *Societal issues,* people perceive the mentioned companies as the most morally unacceptable, especially in terms of being unfair and causing harm to society. In comparison, the category of news articles *Societal Issues-Minorities-Racial Issues* reflects the most morally wrong companies that people are disgusted with. This result highlights the impact of fake news on people's moral perception of a company, which may not align with reality.

How categories of fake news articles differ from trustworthy ones. To measure the extent to which each category i is associated with fake news as opposed to trustworthy news, we calculated the *misinformation association* $MA(i)$ for each article category i: $MA(i) = \frac{P(i|fake)}{P(i|trustworthy)}$. We found that the top 3 categories associated with fake news outlets are *Product-Issues-Quality*, *Politics-Lobbying*, and *Societal Issues-Conspiracy*, while the top 3 associated with trustworthy news outlets are *Brand promotion*, *Corporate affairs-acquisition*, and *Politics-regulation*. This demonstrates that certain categories of news articles, such as product-issues-quality, were more likely to be covered by *fake* news outlets, compared to those *trustworthy* ones.

5 Which Companies Are Targeted?

To answer this question, we identified which companies tend to be targeted, and which factors determined that targeting.

5.1 Companies and Fake News Outlets

Out of the 262 S&P 500 companies discussed on Reddit during the entire period of study T_{all}, 114 were targeted by fake news, and 148 were not. What distinguishes those two types of companies? In Table 6, we show the statistics of our metrics' distributions across the two sets of companies. The two sets of companies statistically differ in their public popularity (`# Reddit comments`, `# Reddit comments per capita` and `# reputable news`) and their `external reputation score`. In that, companies targeted by fake news are more popular and also discussed in more negative tone.

To investigate the specific relationship between companies and fake news outlets targeting them, we generated a bipartite graph of company-outlet interactions in which a fake news outlet and a company are connected to each other if the company is targeted by the outlet. The bipartite graph for the top 50 companies and fake news outlets are shown in Figure 3. We found that: (1) Internet based companies, such as Facebook, Amazon, Microsoft and Apple, attract more fake news. (2) Certain news outlets cover many companies (e.g., `breitbart.com` and `dailywire.com`), whereas other outlets focus on reporting a particular sector (e.g., `stonecoldtruth.com` focuses on communication service companies).

5.2 Factors Associated With the Targeting

For the 126 companies targeted at least once by fake news, we calculated the probability that news discussing them are fake (`pFake` (C, T_{all})) and asked which company characteristics influence this probability. To that end, we ran a linear regression with `pFake` (C, T_{all}) as the dependent variable and the company metrics in Table 3 as independent variables. We confirmed that there is no multicollinearity among our independent variables using Variance Inflation Factors (VIF) (Miles 2014). VIF values above 5 would suggest multicollinearity issues, while the maximal VIF value we found was 1.44, well below that threshold (5). Specifically, we ran

	targeted companies			non-targeted companies				
	mean	median	std	mean	median	std	mean diff	t-stat
Glassdoor rating (C, T_{all})	3.46	3.43	0.35	3.39	3.40	0.33	0.07	1.49
internal employee stress score (C, T_{all})	0.01	0.01	0.00	0.01	0.01	0.01	-0.00	-0.97
stock growth (C)	0.39	0.51	0.93	0.51	0.53	0.56	-0.12	-1.16
log(# **Reddit comments** (C, T_{all}))	9.40	9.31	2.10	5.28	5.52	2.33	**4.12**	14.22^{***}
log(# **Reddit comments per capita** (C, T_{all}))	1.99	1.99	2.02	-1.04	-0.78	2.11	**3.03**	11.25^{***}
external reputation score (C, T_{all})	0.77	0.77	0.01	0.78	0.77	0.08	**-0.02**	-2.13^{**}
log(# **reputable news** (C, T_{all}))	7.40	7.40	1.90	3.02	3.07	1.87	**4.39**	17.82***

Table 6: Differences for company metrics between companies that are targeted by fake news and those that are not. For each of the metrics, we get a set of values, each corresponding to a company. Mean, median and standard deviation, and difference between targeted and non-targeted companies with t-stats are reported. The p-values for the t-stats significance levels are: $.1(*)$, $.05(**)$, and $0.01(***)$. The statistically significant metrics are in bold. All the metrics are calculated within the time period T_{all} (between 2016 and 2019). Compared to non-targeted companies, targeted ones attract more comments on Reddit, have a slightly worse `external reputation`, and significantly more `reputable news` about them shared.

three linear regression models, with the following sets of independent variables:

M1: `log(# Reddit comments per capita)`, `log(# reputable news)`, `Glassdoor rating` and `stock growth`

M2: **M1** + `external reputation score`

M3: **M2** + `internal employee stress score`

The prediction results demonstrated that, despite Model **M1** being the base model, its three predictors explain around 40% of the variance in `pFake` ($Adj.R^2$=.40). Adding the `external reputation score` in **M2** increases the explanatory power to $Adj.R^2$=.42, and also adding the `internal employee stress score` in **M3** leads to an $Adj.R^2$=.43.

Taking the best linear regression model (**M3**), we inspected its β coefficients (Table 7). The coefficient p-values are significant for company popularity (`# Reddit comments per capita`), newsworthiness (`# reputable news`), and external valuations (`stock growth`). Companies that enjoy higher popularity among the general public (on Reddit) and reputable news publishers, yet possess lower stock valuations, tend to be more associated with fake news.

While the coefficients for `external reputation` and `internal employee stress score` appear insignificant, there are increasing scores from $M1$ to $M2$ (when we added `external reputation`), and from $M2$ to $M3$ (when we added `internal employee stress score`). When considering these scores' correlation with the number of reputable news sources (`# reputable news`) and factoring in the results from the subsequent cross-lagged analyses reported in (§6.2), it suggests their association with fake news.

To sum up, companies that suffer severely from fake news are considered newsworthy, garnering more attention from reputable news publishers, yet are perceived negatively both internally and externally. Internally, employees are more likely to comment on a stressful work environment. Externally, investors do not highly value these companies and the general public on social media talks less about them; and if they do mention them, they do so using negative sentiment.

Feature	coef	std error	p-value
Intercept	**-4.06**	0.87	0.00
`log(#Reddit per capita)`	**-4.60**	0.70	0.00
`log(#reputable news)`	**2.80**	0.83	0.00
`stock growth`	**-1.63**	0.85	0.06
`exter. reputation`	-1.48	0.96	0.13
`inter. employee stress`	1.12	0.72	0.13
`Glassdoor rating`	-0.68	0.77	0.38

Table 7: Coefficients and statistics for M3 predicting `pFake` (C, T_{all}) from our company metrics. All metrics are calculated within the time period of T_{all} (between 2016 and 2019). A company that is likely to be targeted by fake news is the subject of more reputable news as well (i.e., it tends to be more newsworthy), and is not doing well in the stock market (lower stock growth). The significant coefficients with p-values at significance level above .1 are shown in bold.

5.3 Factors that Influence the Most Popular Categories of Fake News Articles

To understand the determinants that impact only specific categories of fake news articles (Table 4), we ran three linear regressions having the probability of a fake news article belonging to each of the top 3 categories (*Product*, *Politics*, and *Societal Issues*, as defined in §4.3) as a dependent variable, and the five company metrics of `# Reddit comments per capita`, `internal employee stress score`, `external reputation score`, `stock growth`, and `Glassdoor rating` as independent variables. Results in Table 8 show companies that are mentioned in conjunction with the *Product* and *Societal Issues* attract more comments on Reddit and do well in the stock market as they tend to be tech companies such as Apple. This indicates when covering high-growth companies, fake news outlets tend to cover popular and controversial topics. However, companies linked to fake news in the realm of *Politics* tend to receive lower ratings from their employees. – these typically are pharmaceutical companies such as Pfizer.

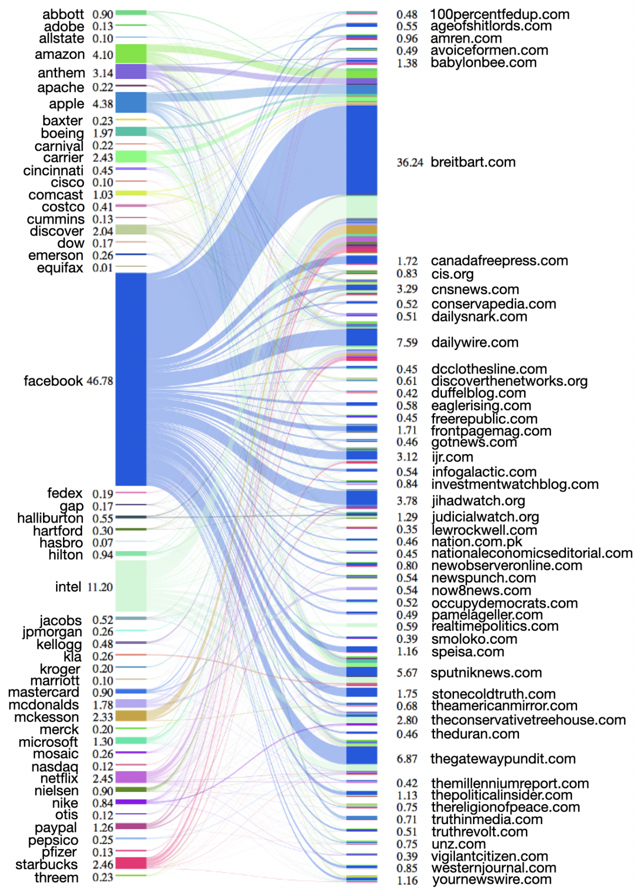

Figure 3: Bipartite graph connecting companies (left hand side) with the fake news outlets (right hand side) mentioning them. A company that is likely to be target by fake news tends to be a tech company such as Facebook and Apple. Tech companies tend to be targeted by multiple outlets, while communication ones (e.g., Google) tend to be targeted by specific outlets such as `stonecoldtruth.com`.

6 When Companies Are Targeted?

Our next research question is to understand *when* companies are targeted by fake news (*RQ3*). We are particularly interested in finding out whether a company is targeted by fake news all the time, or within those concentrated periods, i.e., misinformation shocks (as we defined in §3.5).

Specifically, we identified misinformation peaks for company C by identifying those months that have $pFake(C, T_{month})$ 2 standard deviations above the mean of $pFake(C, T_{all})$ within T_{all}. In Figure 4, we showed several examples of `misinformation shocks` plus its most aligned metric among the three: `stock valuation` (C, T_{month}), `external reputation score` (C, T_{month}) and `internal employee stress score`(C, T_{month}). The companies that are shown are those for which, based on visual inspection, misinformation shocks were aligned the most with at least one of the three metrics. We should stress that such an alignment is a simple association and does not speak to any causal claim though. All the metrics are normalized using z-score throughout the entire periods of study (2016-2019).

Feature	*Product*	*Politics*	*Soc. Issues*
Intercept	-1.22*	-2.14*	-1.87*
# Reddit per capita	2.41*	-0.73	1.63*
external reputation	-0.35	0.72	-0.55
Glassdoor rating	-0.89	-1.24*	-0.76
internal employee stress	0.12	0.24	-0.10
stock growth	1.45*	0.54	1.02*
Adjusted R^2	0.27	0.20	0.16

Table 8: The $Adj.R^2$ and beta coefficients of the linear regression predicting $p(article\_category|fake)$ for three categories of news articles. Companies targeted by fake news attract more comments on Reddit. However, those that are mentioned in conjunction with product and societal issues do well in the stock market as they are tech companies such as Apple, while those that are mentioned in conjunction with politics tend to do worse among their employees (in terms of their company review ratings) and typically are pharmaceutical companies such as Pfizer. Significant values are marked with * based on their significance levels $p < 0.05$.

6.1 Qualitative Case Studies

Apple. The company experienced two distinct instances of misinformation, or "misinformation shocks", in late 2016 and early 2017. The majority of the articles during these periods focused on three main topics: the FBI's alleged request for Apple to add "back-doors" in their phones; false claims of a potential recall of their phones due to alleged issues with them catching fire; and a ransomware attempt in which hackers threatened to wipe out 300 million iPhone accounts in early 2017, however, there was no evidence that the company was affected by this attempt. After the second misinformation shock, the company's `stock valuation` decreased by 0.2 standard deviations.

Google. The company experienced one instance of misinformation, or "misinformation shock", in late 2016. The articles during this period focused on a single claim that the company's ranking algorithm downgraded any news stories that would harm the then presidential nominee of the Democratic party, Hillary Clinton. This claim was unproven and categorically rejected by the company. As a result of this misinformation shock, the company's *external reputation score* decreased by 1.4 standard deviations.

Facebook. The company experienced two instances of "misinformation shocks". The first occurred in early 2016 and centered around the claim that the company was not complying with European laws requiring the removal of hate speech on its platform. One month after this shock, the company's `internal stress score` increased by 2.3 standard deviations. The second shock occurred in early 2017 and involved two separate events related to discrimination. The first event was an allegation that the platform had an internal document advising on how to censor news of crimes against Black people. The company acknowledged this document was no longer in use. The second event was an allegation

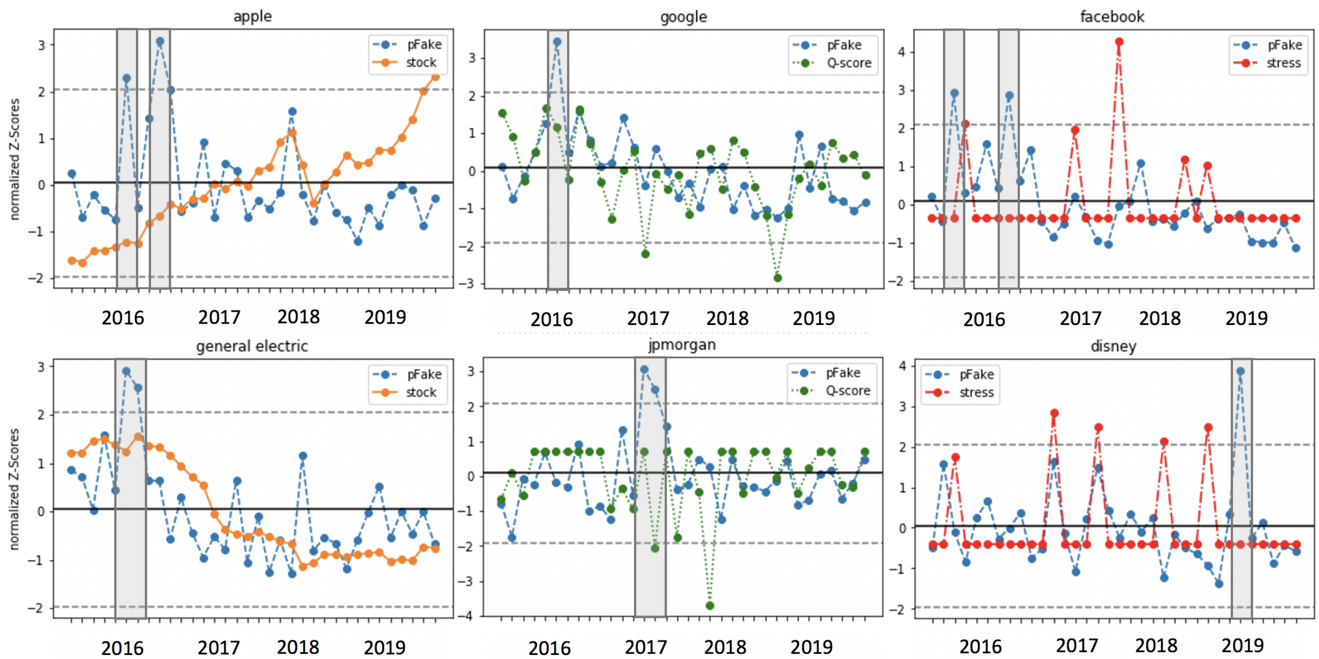

Figure 4: Fake news shocks (peaks of $pFake(C, T_{month})$ over time, shown with the gray bars) for a company C that: is a Tech company (Apple and Google); is a Communication company (Facebook); has not done well in the stock market (General Electric); attracts considerable negative mentions on social media (JPMorgan); and attracts considerable stress mentions by its employees (Disney).

that the company's senior engineers blocked diversity hires. The company declined to comment on this claim but stated that they "care deeply about diversity".

General Electric. The company experienced a single instance of significant "misinformation shock" that lasted for two months in late 2016. The majority of the articles during this period focused on a rumor that the company intended to sell its $3 billion Industrial Solutions business. After the circulation of this rumor, the company's `stock valuation` decreased by 0.2 standard deviations. It was later revealed the rumor was true, as the sale of the business was initiated in late 2017 and completed in mid-2018.

JP Morgan. The company experienced a single instance of significant "misinformation shock" that lasted for two months in late 2017. The majority of the articles during this period focused on a scandal in which the company, JP Morgan, was alleged to have transferred $875 million into a former Nigerian oil minister's account, and was subsequently sued by the Nigerian government. After this, the company's `external reputation score` decreased by 2.7 standard deviations. However, in 2022, a judge at the British High Court ruled in favor of JP Morgan.

Disney. The company experienced a single instance of significant "misinformation shock" in mid-2019. The majority of the articles during this period focused on allegations of child pornography involving former or current employees. In one instance, a former employee was arrested and charged with "transferring obscene materials to a minor". This shock did not have a significant impact on the company's *internal stress score*, unlike other minor instances of misinformation that had occurred in the past, such as slow progress in building Disney's Pandora land in 2017 and allegations of CIA aid in the company's land acquisition in Florida in 2018, which did.

These are just examples of a general trend that we observed: a company is not targeted by fake news all the time, but there are particular times in which a critical mass of fake news emerges, and that might impact either the company's *external reputation*, *internal employee stress* or even *stock valuation*.

6.2 Cross-lagged Analysis

To further examine how fake news impact company's external reputation, internal employee stress and stock valuation, we conducted cross-lagged analysis. The goal is to determine causal relationships by examining the connection between two variables at different points in time, while taking into account the influence of earlier measurements of the same variables. This type of analysis helps eliminate the possibility that the observed relationship between the variables is caused by a third or confounding variable. In our study, we conducted cross-lagged analysis on all time pairs (t_k, t_l) and selected the pair and the model based on three conditions. The selected cross-lagged model must: 1) demonstrate significantly stronger effects compared to a constrained model (Kenny 1975), as determined by an ANOVA test; 2) have cross-lagged effects that are statistically significant (at least $p < 0.05$); and 3) have the highest absolute mean cross-lagged effect among all the other models. To interpret the values of cross-lagged effect, we employ Cohen's statistical

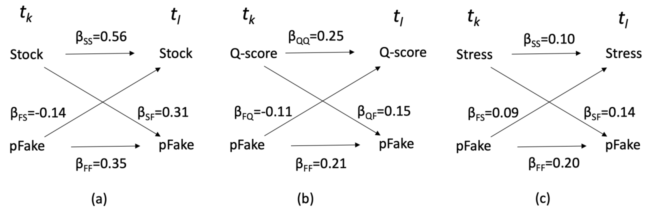

Figure 5: Cross-lagged analysis between p**F**ake and various variables: (a) stock valuation (**S**tock), (b) external reputation (**Q**-score) and (c) internal employee stress (**S**tress).

effect size (Cohen 2013), where an effect size smaller than 0.2 is considered weak, an effect size between 0.2 and 0.5 is considered moderate, and an effect size between 0.5 and 1.0 is considered strong. The cross-lagged analysis results (Figure 5) indicate that:

Stock valuation (Figure 5a): we found evidence when a company's stock value goes up, it tends to continue going up in the future. This was shown by a statistical analysis that gave a result of $\beta_{SS} = .56$ with a probability of less than 1% that it occurred by chance ($p < 0.01$). Furthermore, we noticed a moderate tendency for fake news probability (pFake) to reinforce itself ($\beta_{FF} = .35$, $p < 0.01$). This implies that when there is a high likelihood of fake news, it is likely to be even higher in the future. In addition, we looked at the relationship between a company's stock value at time t_k and the occurrence of fake news about that company at a later point in time t_l. The findings showed that there is a moderate, positive connection between the two (standardized $\beta_{SF} = .31$, $p < 0.01$). This suggests that when a company's stock value is higher at an earlier point in time, it increases the probability that the company will be the target of fake news at a later point in time. In contrast, we found that exposure to fake news at time t_k had a weak and negative effect (standardized $\beta_{FS} = -.14$, $p < 0.05$) on a company's stock value at a later time t_l. This means fake news has a detrimental effect (albeit small) on a company's valuation, potentially leading to a decrease in its stock value.

External reputation (Q-score) (Figure 5b): moderate evidence indicated a positive self-reinforcement of external reputation ($\beta_{QQ} = .25$, $p < 0.01$) and pFake ($\beta_{FF} = .21$, $p < 0.01$), suggesting that higher levels of external reputation and pFake were associated with subsequent increases in these variables. Regarding the cross-lagged effects, we examined the influence of Q-score at t_k on fake news occurrence at t_l, where t_l follows t_k. Results indicated a weak positive effect ($\beta_{QF} = .15$, $p < 0.05$), suggesting that higher Q-scores at an earlier time slightly increased the likelihood of future fake news targeting. Conversely, a weak negative effect ($\beta_{FQ} = -.11$, $p < 0.05$) was observed, indicating that fake news exposure at t_k had a small negative impact on subsequent external reputation.

Internal employee stress (Stress) (Figure 5c): there was weak evidence of stress self-reinforcement ($\beta_{SS} = .10$, $p < 0.10$), indicating a slight association between higher stress levels and subsequent stress increases. Moderate self-reinforcement of pFake was found ($\beta_{FF} = .20$, $p < 0.01$), linking higher pFake levels to increased pFake in the future. Regarding the cross-lagged effects, stress at t_k positively influenced the likelihood of future fake news targeting at t_l with a weak magnitude ($\beta_{SF} = .14$, $p < 0.05$). Similarly, fake news exposure at t_k had a weak positive effect on subsequent stress levels at $t_l(\beta_{FS} = .09, p < 0.10)$.

To summarize, our cross-lagged analysis revealed that higher past stock valuations increase the future risk of fake news targeting. Weaker effects were observed for external reputation, where higher past reputation makes companies prone to future fake news targeting. Conversely, being targeted by past fake news leads to reputation decline. Regarding internal stress, high past stress levels make companies susceptible to future fake news targeting, and past targeting increases future employee stress. These findings highlight the interplay between stock valuation, reputation, stress, and fake news, indicating risks for targeted companies.

7 Discussion

Limitations. Our study has five limitations. First, we attribute fake news to outlets rather than individual articles. Nonetheless, research indicates that publisher-level credibility scores align closely with article-level ones (Xu 2021). Second, our analysis of company perceptions is based on data from Reddit and Glassdoor, so it is unclear if the conclusions hold true on other platforms. Specifically, Reddit is an anonymous platform without the concept of 'friends', unlike many other social networks. As such, this platform is missing social pressure, and users are less likely to enter echo chambers, which means that misinformation spreading is not as likely to stem from homophily (McPherson, Smith-Lovin, and Cook 2001) at the circle-of-friends level as it is on other social networks. Third, we studied the sentiment of news mentioning a company and stress mentions in company reviews written by employees. These aspects are often overlooked. We used state-of-the-art methods to proxy these metrics, and our results show that they are partly predictive of misinformation. However, our study does not capture all important factors that influence public perception of a company, and might not generalize for other industries or countries. For example, we did not consider corporate social responsibility (Lindgreen and Swaen 2010) or brand identity (Balmer 2008). These aspects are still subject of ongoing research, and we leave their exploration for future work. Fourth, our data did not include comments that were deleted before being collected, so we could not examine if they contained news URLs. In particular, comments that were removed by Automoderator (bots) were unavailable to us, as these comments were removed as soon as they were posted. Nevertheless, the Reddit dataset from `pushshift.io` remains one of the most comprehensive datasets available (Baumgartner et al. 2020). Fifth, our study was conducted on a sample of S&P 500 companies in the United States. It is possible that our findings may not generalize to other industries, countries, or time periods. Future research is needed to further investigate the generalizability of our findings.

Implications. Our study has three main *theoretical implications*. First, we provide a fine-grained taxonomy of fake news coverage in the corporate world. Second, we identify the factors that contribute to a company's vulnerability to misinformation (e.g., a company's external and internal perceptions). Third, we contribute to understanding the temporal patterns of the misinformation, and how it varies over time. Our findings offer an initial step towards the development of a theoretical framework for understanding the complex phenomenon of corporate misinformation.

Our study has four main *practical implications*. First, our findings help in identifying which companies are more susceptible to misinformation such as financial performance and reputation, allowing them to take more proactive measures to prevent the spread of false information. Second, understanding the categories of corporate misinformation that are prevalent, allows companies and stakeholders to better recognize and develop tools and strategies for combating corporate misinformation. Third, by tracking temporal patterns of misinformation, and detecting "misinformation shocks" early on, companies can anticipate and promptly respond to those potential misinformation crises. Finally, our results demonstrate the importance of a systematic monitoring system for companies, as well as transparent communication with employees to manage a possible misinformation crisis better and faster (Holtom, Edmondson, and Niu 2020).

Broader Perspective, Ethics and Competing Interests. Our research originates from a lab within a for-profit, traded company. As such, the authors may benefit financially from the implementation of stronger measures against perpetrators of misinformation. We acknowledge this potential conflict of interest and affirm that the research was conducted with integrity and that the results presented are unbiased and fully reflect all our findings. Our datasets were collected from publicly available sources. We also confirm that our conclusions are based solely on the data and analysis presented, and are not influenced by any financial gain or benefit that may result from the implementation of these measures. We have used rigorous statistical methods to analyze the data and drawn conclusions that are supported by evidence. Our study was reviewed by the internal company ethical committee while the study protocol, data collection methods, informed consent process, and other aspects of the study ensured that we met ethical and legal standards. The part of the study that involved human subjects was a crowdsourcing user study, where participants were paid at least the minimum wage. All our analysis were based on aggregated data without tracking down to individuals.

8 Conclusion

Our study delved into the complex issue of corporate misinformation, specifically focusing on the S&P 500 companies. Utilizing a combination of large-scale social media and crowd-sourcing data, we uncovered key insights into the nature of fake news targeting these companies. We found that fake news outlets often report on a company's products, politics, and societal issues when spreading misinformation. Furthermore, we established a clear link between a company's external and internal perceptions and its likelihood of being targeted by fake news. Companies that are struggling in the stock market, receiving negative mentions on social media, and whose employees are facing stress are more likely to be targeted by fake news. Additionally, we discovered that fake news about a company does not occur consistently, but rather emerges in specific, concentrated periods. These findings have significant implications for companies and stakeholders looking to combat the spread of misinformation.

References

2021. GameStop Shares Surge 100% In A Day, Reddit Group Rejoices.

Adriani, R. 2019. Fake news in the corporate world: A rising threat. *European Journal of Social Science Education and Research*, 6(1): 92–110.

Ahmed, W.; Vidal-Alaball, J.; Downing, J.; Seguí, F. L.; et al. 2020. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. *Journal of medical internet research*, 22(5): e19458.

Akbik, A.; Bergmann, T.; Blythe, D.; Rasul, K.; Schweter, S.; and Vollgraf, R. 2019. FLAIR: An easy-to-use framework for state-of-the-art NLP. In *NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, 54–59.

Al-Rawi, A.; and Fakida, A. 2021. The methodological challenges of studying "fake news". *Journalism Practice*, 1–20.

Alexa.com. 2022. Alexa sites.

Allcott, H.; and Gentzkow, M. 2017. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2): 211–36.

Allcott, H.; Gentzkow, M.; and Yu, C. 2019. Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2): 2053168019848554.

Balmer, J. M. 2008. Identity based views of the corporation: Insights from corporate identity, organisational identity, social identity, visual identity, corporate brand identity and corporate image. *European journal of marketing*, 42(9/10): 879–906.

Balsamo, D.; Bajardi, P.; and Panisson, A. 2019. Firsthand opiates abuse on social media: monitoring geospatial patterns of interest through a digital cohort. In *The World Wide Web Conference*, 2572–2579.

Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, 830–839.

Bozarth, L.; Saraf, A.; and Budak, C. 2020. Higher ground? How groundtruth labeling impacts our understanding of fake news about the 2016 US presidential nominees. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 48–59.

Bronnenberg, B. J.; Dubé, J.-P.; and Sanders, R. E. 2020. Consumer misinformation and the brand premium: A private label blind taste test. *Marketing Science*, 39(2): 382–406.

Camacho, M. M.; and Lozano, A. 2020. Disinformation as a corporate risk: the need for a new methodological approach. *Vivat Academia*, (155): 111–129.

Castellani, P.; and Berton, M. 2017. Fake news and corporate reputation: What strategies do companies adopt against false information in the media? In *Proc. Toulon-Verona Conference" Excellence in Services*.

Cheng, Y.; and Chen, Z. F. 2020. The influence of presumed fake news influence: Examining public support for corporate corrective response, media literacy interventions, and governmental regulation. *Mass Communication and Society*, 23(5): 705–729.

Chung, W.; Zhang, Y.; and Pan, J. 2022. A Theory-based Deep-Learning Approach to Detecting Disinformation in Financial Social Media. *Information Systems Frontiers*, 1–20.

Cohen, J. 2013. *Statistical power analysis for the behavioral sciences*. Academic press.

Cover, R.; Haw, A.; and Thompson, J. D. 2022. Marginalising the Marginalised: Fake News as a Tool of Populist Power. In *Fake News in Digital Cultures: Technology, Populism and Digital Misinformation*. Emerald Publishing Limited.

Di Domenico, G.; and Visentin, M. 2020. Fake news or true lies? Reflections about problematic contents in marketing. *International Journal of Market Research*, 62(4): 409–417.

Dot, D. 2022. Fake News Sites list Facebook.

Eady, G.; Paskhalis, T.; Zilinsky, J.; Bonneau, R.; Nagler, J.; and Tucker, J. A. 2023. Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nature Communications*, 14(1): 62.

Edge, L. 2021. Scientific misinformation: A perfect storm, missteps, and moving forward. *Cell*, 184.

Fox, J. 2020. 'Fake news'–the perfect storm: historical perspectives. *Historical Research*, 93(259): 172–187.

Freeman, K. D.; and Dart, J. 1993. Measuring the perceived quality of professional business services. *Journal of Professional Services Marketing*, 9(1): 27–47.

Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2018. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the world wide web conference*, 913–922.

Gillin, J. 2018. PolitiFact's guide to fake news websites and what they peddle. Politifact.

Graham, J.; Nosek, B. A.; Haidt, J.; Iyer, R.; Koleva, S.; and Ditto, P. H. 2011. Mapping the moral domain. *Journal of personality and social psychology*, 101(2): 366.

Holtom, B.; Edmondson, A.; and Niu, D. 2020. Tips for communicating with employees during a crisis. *Harvard Business Review*.

Jahng, M. R. 2021. Is fake news the new social media crisis? Examining the public evaluation of crisis management for corporate organizations targeted in fake news. *International Journal of Strategic Communication*, 15(1): 18–36.

Kalsnes, B. 2018. Fake news. In *Oxford Research Encyclopedia of Communication*.

Kenny, D. A. 1975. Cross-lagged panel correlation: A test for spuriousness. *Psychological bulletin*, 82(6): 887.

Kıcıman, E.; Ness, R.; Sharma, A.; and Tan, C. 2023. Causal reasoning and large language models: Opening a new frontier for causality. *arXiv preprint arXiv:2305.00050*.

Kogan, S.; Moskowitz, T. J.; and Niessner, M. 2019. Fake news: Evidence from financial markets. *SSRN*, 3237763.

Lazer, D. M.; Baum, M. A.; Benkler, Y.; Berinsky, A. J.; Greenhill, K. M.; Menczer, F.; Metzger, M. J.; Nyhan, B.; Pennycook, G.; Rothschild, D.; et al. 2018. The science of fake news. *Science*, 359(6380): 1094–1096.

Lin, T. C. 2016. The new market manipulation. *Emory LJ*, 66: 1253.

Lindgreen, A.; and Swaen, V. 2010. Corporate social responsibility. *International journal of management reviews*, 12(1): 1–7.

MBSFC. 2022. Media Bias Fact Check.

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 415–444.

Melchior, C.; and Oliveira, M. 2022. Health-related fake news on social media platforms: A systematic literature review. *new media & society*, 24(6): 1500–1522.

Miles, J. 2014. Tolerance and variance inflation factor. *Wiley statsref: statistics reference online*.

Mills, A. J.; and Robson, K. 2019. Brand management in the era of fake news: narrative response as a strategy to insulate brand value. *Journal of Product & Brand Management*.

Moshood, T. D.; Shittu, R. A.; and Abidin, T. 2020. Covid-19 and 5G radiation are two parallel lines: A systematic review. *Int. J. Innov. Sci. Res. Technol*, 5: 744–751.

Muzykant, V. L.; Muqsith, M. A.; Pratomo, R. R.; and Barabash, V. 2021. Fake news on COVID-19 in Indonesia. In *Pandemic Communication and Resilience*, 363–378.

Naeem, S. B.; Bhatti, R.; and Khan, A. 2021. An exploration of how fake news is taking over social media and putting public health at risk. *Health Information & Libraries Journal*, 38(2): 143–149.

Ng, L. H.; and Taeihagh, A. 2021. How does fake news spread? Understanding pathways of disinformation spread through APIs. *Policy & Internet*, 13(4): 560–585.

Park, A.; Montecchi, M.; Plangger, K.; Pitt, L.; et al. 2020. Understanding fake news: a bibliographic perspective. *Defence Strategic Communications*, 8(Spring 2020): 141–172.

Pennycook, G.; and Rand, D. G. 2021. The psychology of fake news. *Trends in cognitive sciences*, 25(5): 388–402.

Pickard, V. 2019. The Misinformation Society. In *Antidemocracy in America*, 39–48. Columbia University Press.

PWC. 2022. Corporate Sector Disinformation.

Scepanovic, S.; Martin-Lopez, E.; Quercia, D.; and Baykaner, K. 2020. Extracting medical entities from social media. In *Proceedings of the ACM Conference on Health, Inference, and Learning*, 170–181.

Serazio, M. 2021. The other 'fake'news: Professional ideals and objectivity ambitions in brand journalism. *Journalism*, 22(6): 1340–1356.

Starbird, K. 2021. Online Rumors, Misinformation and Disinformation: The Perfect Storm of COVID-19 and Election 2020.

Swire-Thompson, B.; and Lazer, D. 2019. Public health and online misinformation: challenges and recommendations. *Annual review of public health*, 41: 433–451.

Tandoc Jr, E. C. 2019. The facts of fake news: A research review. *Sociology Compass*, 13(9): e12724.

Vargo, C. J.; Guo, L.; and Amazeen, M. A. 2018. The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New media & society*, 20(5): 2028–2049.

Verma, G.; Bhardwaj, A.; Aledavood, T.; De Choudhury, M.; and Kumar, S. 2022. Examining the impact of sharing COVID-19 misinformation online on mental health. *Scientific Reports*, 12(1): 1–9.

Vosoughi, S.; Roy, D.; and Aral, S. 2018. The spread of true and false news online. *science*, 359(6380): 1146–1151.

Waisbord, S. 2018. Truth is what happens to news: On journalism, fake news, and post-truth. *Journalism studies*, 19(13): 1866–1878.

Westerlund, M. 2019. The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11).

Wikipedia. 2022. S&P 500 Companies.

Xu, R. 2021. *Corporate Fake News on Social Media*. Ph.D. thesis, University of Miami.

Yahoo Finance portal. 2022. S&P 500 Stock Market Data.

Zhou, X.; and Zafarani, R. 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, 53(5): 1–40.

Zimdars, M. 2016. My "fake news list" went viral. But made-up stories are only part of the problem. *Washington Post*.

A Ethics Checklist

1. Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes
2. Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes
3. Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes
4. Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes
5. Did you describe the limitations of your work? Yes
6. Did you discuss any potential negative societal impacts of your work? Yes
7. Did you discuss any potential misuse of your work? Yes
8. Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes
9. Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes
10. Did you include the full text of instructions given to participants and screenshots? Yes
11. Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? Yes
12. Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? Yes
13. Did you discuss how data is stored, shared, and deidentified? Yes

B Appendix

Our paper has five main findings[1]:

1. Companies that suffer severely from fake news are also likely to be covered by reputable news publishers.
2. Companies susceptible to fake news receive less attention from the general public on social media.
3. Being targeted by fake news leads to a higher likelihood of public comments with negative sentiments on social media.
4. A company targeted by fake news is likely to experience a lower stock market growth.
5. In companies targeted by fake news, employees are more likely to express stress indicators in their reviews.

In recent advancements, the efficacy of Large Language Models (LLMs) in elucidating causal relationships has emerged as a potential application (Kıcıman et al. 2023). Building upon this work (Kıcıman et al. 2023), we embark on an exploration of causal analysis leveraging the capabilities of GPT4, a state-of-the-art language model[2]. Our primary objective revolves around validating our five findings. To achieve this, we adopt prompt strategies from (Kıcıman et al. 2023) aimed at validating each causal claim by evaluating its normality or abnormality. A typical prompt includes:

[1] https://social-dynamics.net/fake-news/

System: You are a helpful assistant for causal reasoning. Specifically, you are an advanced expert specializing in causal reasoning within the realm of Corporate Misinformation research using public data, such as stock price, published news, and social media responses. Equipped with a strong knowledge base and experience in financial, social and environmental determinants of misinformation in the company world, you are proficient in identifying, assessing, and interpreting the complex interactions among variables. Your primary responsibility involves validating known or hypothesized causal relationships between a diverse array of variables and their respective outcomes.

User: You will receive a claim on the cause-and-effect relationship between variables. Based on all the information, given a causal event that causes an outcome event at a later time point, the causal event is "abnormal" if the causal event's occurrence was unexpected, unlikely, surprising, rare, or improbable. The opposite of "abnormal" is "normal". A causal event is 'normal' if its occurrence was expected, likely, probable, and not surprising. The normality of a causal event is independent of whether the outcome event was bad or and whether the outcome event was normal. Please read the following statement about a causal event. Provide your step-by step rationale as to why the causal event is abnormal or normal. Then, return the word "normal" if you think the causal event is normal or "abnormal" if you think the causal event is abnormal.

Claim: *Companies that suffer severely from fake news are also likely to be covered by reputable news publishers.*

The step-by-step answer of GPT-4 to that typical prompt unfolds as follows: "The claim suggests a causal relationship between severe suffering from fake news by companies and increased coverage by reputable news publishers. Let's break this down step by step:

Companies suffering severely from fake news: This implies a situation where a company is significantly affected by false information spread about it.

Likelihood of coverage by reputable news publishers: The claim implies that when a company faces severe fake news attacks, reputable news outlets are more likely to cover the situation.

Rationale for Abnormality or Normality:

- *Abnormal*: This relationship might be considered abnormal because, traditionally, reputable news publishers tend to verify and fact-check information rigorously before reporting. Therefore, if a company is suffering from fake news, the expectation might be that reputable news sources would be cautious about covering unverified or

[2] https://openai.com/blog/gpt-4-api-general-availability

misleading information, potentially leading to less coverage rather than more.

- *Normal*: On the other hand, in today's media landscape, there's a growing emphasis on reporting on significant events, even if they are controversial or involve misinformation. Companies facing severe fake news might attract attention due to the sensational nature of the situation, prompting reputable news outlets to cover it to clarify the facts and counteract the false information.

Considering the dynamics of sensationalism in media and the role of reputable news sources in countering fake news, the relationship might indeed be normal. While it might seem unexpected for reputable sources to cover such companies more, the desire to address and rectify false information could drive increased coverage. *Therefore, the claim is* normal *in this scenario.*

The resulting assessments of the (ab)normality of our five claims (Table 9) suggest that all five of our main findings are supported. However, the LLM does not always return 'normal' as an answer. To illustrate this, we also considered statements for which our causal analysis did not have strong evidence to be conclusive. One such statement is reported in the last row of Table 9 and suggests that companies in which employees express stress indicators when reviewing their workplaces are more likely to be targeted by fake news. Indeed, the LLM found this claim to be abnormal.

Statement	Conclusion
Companies that suffer severely from fake news are also likely to be covered more by reputable news publishers.	Considering the dynamics of sensationalism in media and the role of reputable news sources in countering fake news, the relationship might indeed be normal. While it might seem unexpected for reputable sources to cover such companies more, the desire to address and rectify false information could drive increased coverage. *Therefore, the claim is **normal** in this scenario.*
Companies susceptible to fake news receive less attention from the general public on social media.	Considering the dynamics of how fake news spreads on social media and the potential hesitancy to engage with content that might be false, the relationship between companies being subject to fake news and the general public talking less about them on social media might be considered normal. The decrease in discussion could be a result of cautious behavior among the public to avoid amplifying false information. *Therefore, the claim is **normal** in this scenario.*
Being targeted by fake news leads to a higher likelihood of public comments with negative sentiments on social media, contributing to the decline in the reputation of affected companies.	Considering the typical response patterns to fake news and its impact on reputation, the relationship between being targeted by fake news, experiencing reputation decline, and receiving negative comments on social media is more likely to be normal. The emotional response triggered by fake news and its effect on public sentiment could drive increased negative comments on social media. *Therefore, the claim is **normal** in this scenario.*
A company targeted by fake news is likely to experience a lower stock market growth.	Considering the complexities of the stock market and the multitude of factors influencing stock prices, the relationship between being likely to be targeted by fake news and maintaining lower stock growth might be normal. While other market dynamics are at play, the impact of negative news or misinformation on stock performance could lead to lower growth. *Therefore, the claim is **normal** in this scenario.*
In companies targeted by fake news, employees are more likely to express stress indicators in their reviews.	Considering the potential impact of negative external events on employee morale and the prevalence of stress-related discussions in workplace environments, the relationship between companies susceptible to fake news targeting and considerable stress mentions by their employees might be considered normal. The spread of false information about the company could indeed impact the internal environment and employee well-being. *Therefore, the claim is **normal** in this scenario.*
Employees express stress indicators when reviewing their workplaces are more likely to be targeted by fake news.	Considering the link between employee stress indicators and potential targeting by fake news, it might be considered abnormal as the direct causal relationship between these variables could be unexpected or unlikely. The connection might involve other complex factors beyond just stress indicators. *Therefore, the causal event is **abnormal** in this scenario.*

Table 9: Validation of our five main findings and one unsubstantiated claim by prompting LLM GPT-4 for causality as per (Kıcıman et al. 2023). Causal relationships are considered validated if found to be normal.