Transmission Bias in Financial News

Khaled Obaid and Kuntara Pukthuanthong

July 31, 2024

Abstract

We examine how financial news is distorted as it spreads across different news outlets, akin to the "telephone game". Using a sample of exclusive articles from the *Wall Street Journal*, we use ChatGPT-4 to quantify transmission bias as competing news outlets retell these stories. We find strong evidence that retelling articles tend to be more opinionated and negative and less factual and appealing compared to the original story. Less specialized news outlets, the more time-lapse between the original story and its retelling, and the presence of competing narratives from other news outlets contribute to the distortion. Media distortion negatively predicts the following month's abnormal returns, suggesting that transmission bias creates a stronger and more pessimistic public perception of the original story. Transmission bias mostly influences retail traders, and leads to heightened disagreement among sophisticated and unsophisticated traders.

JEL classification: G4, D8, M31

Keywords: Financial news, Telephone game, Distortion, Media bias, Transmission bias, Investor attention, ChatGPT, Large language models.

Khaled Obaid is an Assistant Professor at the College of Business, Mississippi State University, Mississippi State, MS 39762; email kobaid@business.msstate.edu. Kuntara Pukthuanthong is the Robert J. Trulaske Jr. Professor of Finance at the Robert J. Trulaske Sr. College of Business, University of Missouri, Columbia, MO 65211; email pukthuanthongk@missouri.edu.

1. Introduction

In an era dominated by the rapid dissemination of information across many media platforms, financial news could be distorted as it permeates the media ecosystem. In this study, we utilize ChatGPT-4 to evaluate how financial content might be distorted as it spreads across news outlets (i.e., transmission bias) and the financial implications of this distortion. We leverage a sample of exclusive news articles from the *Wall Street Journal* (WSJ) and track how these stories are retold in other news outlets. Oftentimes, investors learn about financial market developments from second or third-hand sources. For instance, the WSJ reported exclusively on June 11, 2021, that Mudrick Capital Management LP experienced a 10% loss in its flagship fund due to an unexpected surge in AMC Entertainment Holdings' stock price.[1] This same story was retold by the *New York Post* (NYP) on June 14, 2021.[2] Some investors will learn about this story from the NYP instead of the WSJ.

The transmission of the original story in retelling articles could potentially be biased. Several factors can drive the bias, including memory (e.g., unusual or unexpected details are more easily remembered), social and motivational factors (e.g., competition), and the specialization of news outlets (Barrett and Nyhof, 2001; Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006). For example, given that negative stories tend to attract more attention, it is possible that news outlets retelling the exclusive WSJ will add an additional negative tone to the story in order to capture the attention of readers. In the example articles above, the percentage of negative words in the original WSJ article and the retelling according to the Loughran and McDonald (2011) lexicon is 2.80% and 5.03%, respectively. Another WSJ exclusive story reported that PepsiCo Inc. is implementing layoffs at the headquarters of its North American snacks and beverages divisions on December 5, 2022.[3] The *Washington Post* (WP) retold this same story on December 6, 2022.[4] Similarly, the

[1]Source: AMC Bet by Hedge Fund Unravels Thanks to Meme-Stock Traders
[2]Source: Mudrick Capital takes massive hit on AMC shares as short bet backfires
[3]Source: PepsiCo to Lay Off Hundreds of Workers in Headquarters Roles
[4]Source: PepsiCo layoffs could indicate broader economic woes

original story has nearly half the proportion of negative words compared to the retelling (1.20% and 2.48%, respectively).

In light of these patterns, we draw parallels to the "telephone game" where the message erodes with each successive retelling. More specifically, we examine the following research questions: (1) How are the original news stories altered when retold (e.g., language style and level of detail)? (2) What news outlet and story characteristics influence this distortion? (3) Can advanced large language models help us quantify how news is distorted as it is retold across news outlets? and (4) Does distortion in retelling articles have implications on asset prices, trading behavior, and the information environment?

We focus our analysis on the full text of exclusive WSJ articles as our starting point. These stories are labeled "exclusive" by the WSJ, indicating that the WSJ has information, an interview, or a story that no other media outlet has at the time of publication or broadcast.[5] This sample serves as our starting point. Next, we create a sample of retelling articles, which is composed of articles from other major news outlets that retell the same story first released by the WSJ. The source of retelling articles comes from major US news outlets in Factiva, and we require articles to directly mention WSJ as the source of the story. We utilize Term Frequency-Inverse Document Frequency (TF-IDF) and cosine similarity to find potential matches and confirm the matches manually.[6]

Between 2013 and 2022, we identify 1,351 exclusive WSJ articles that we are able to match to at least one article from one of the major US news outlets in Factiva. On average, we find that each article is retold 1.40×. Our sample covers retelling articles from 18 news organizations, including the New York Times, Washington Post, New York Post, Investor's

[5]Relying on a sample of general news and using the publication date to determine the story's source is problematic for our study focusing on transmission bias. While the publication date can indicate which news outlet was the quickest to release the story, it may not necessarily be the original source that other news outlets depend on.

[6]Term Frequency-Inverse Document Frequency is a statistic that gauges a word's significance within a document in a corpus, adjusting for its commonness across documents. Cosine similarity measures the similarity between two documents, regardless of size, by calculating the cosine of the angle between their vector representations in multidimensional space. In our study, we apply both methods to identify potential matches.

Business Daily, and USA Today, among others.[7]

After creating our sample of exclusive WSJ articles and respective retelling articles, we utilize ChatGPT-4 to examine how the retelling differs from the original story in five dimensions of language style: Fact, Opinion, Negativity, Positivity, and Appeal. ChatGPT's advantage lies in its ability to understand and interpret complex text in a way that mimics human analysis. Unlike sentiment or topics that can be captured using the dictionary approach, capturing how two articles about the same story differ requires reasoning and attention to context that traditional textual analysis cannot capture. Unlike traditional textual analysis or machine learning approaches, ChatGPT is especially adept at reducing the dimensionality of complex textual data (e.g., Chen, Kelly, and Xiu, 2023). This makes it particularly suitable for analyzing the distortion of financial narratives as they spread across news outlets.

Our first empirical test examines whether there is any bias in the transmission of news. We document that retelling articles are deemed by ChatGPT-4 to be more opinionated and negative and less factual, positive, and appealing relative to the original story. We refer to the shift in writing as negative personalization. Retelling articles receive a negative personalization score that is, on average, 6.99% higher than the original articles. Moreover, we help validate our ChatGPT-4 variables by showing that automated text variables that capture tone (Loughran and McDonald, 2011), readability, and presence of numbers in the text relate to the ChatGPT-4 variables in the predicted direction. We control for these automated text variables in all our tests and show that the ChatGPT-4 language variables contain information beyond what these automated text variables contain.

Next, we explore the drivers behind the negative personalization in retelling articles. Specifically, we consider (1) the frequency at which news outlets retell stories, (2) the time-lapse between the retelling and the original story, and (3) the competition for retelling the story (measured by the number of retelling articles by different news outlets of the same WSJ

[7]Please see Appendix A for the complete list.

story on the same day). We show that retelling articles that are retold later and retold by less specialized news outlets (i.e., news outlets that frequently retell stories) tend to be more opinionated and less factual. Moreover, we show that retelling articles that are retold by more specialized news outlets, published later, and retold in a more competitive environment tend to be more pessimistic. Stories that are retold by less specialized news outlets and in less competitive environments are deemed to be less appealing. Overall, we document clear patterns of negative personalization (i.e., the text becomes more opinionated, more negative, less factual, and less appealing) as news spreads and explore some of the factors that influence the distortion.

Next, we examine whether transmission bias impacts asset prices. Past studies document how news coverage catches investors' attention (e.g., Engelberg and Parsons, 2011; Solomon, Soltes, and Sosyura, 2014) and leads to temporary mispricing driven by biases such as overreaction and stale information (e.g., Dougal, Engelberg, Garcia, and Parsons, 2012; Engelberg, Sasseville, and Williams, 2012; Gurun and Butler, 2012). We find that transmission bias negatively predicts the following month's abnormal returns.[8] Our evidence on predictability is consistent with negative personalization in retelling articles, leading to a more pessimistic public perception of the original story, subsequently resulting in decreased abnormal returns the following month. The negative relation with future abnormal returns is mostly driven by the personalization of the story as opposed to the increase in the negative tone or decrease in the appeal. The predictive effect of retelling articles is economically significant; a retelling article is linked to a 1.07% reduction in the stock return of the featured firm in the following month, which is substantial compared to the mean next-month abnormal return of -0.20%. We later show that this predictability result reverses if we extend the predictability window to 2 or 3 months.

Our evidence is consistent with prior work documenting that investors react to stale in-

[8]Our abnormal returns ($abret_{i,j,t,t+21}$) is the compounded abnormal returns between $t+1$ and $t+21$ adjusted using the Fama-French three-factor model plus momentum for stock j mentioned in article i. We use 100 days estimation window 50 days prior to time t and require a minimum of 70 valid returns to estimate the expected returns.

formation, leading to mispricing (e.g., Huberman and Regev, 2001; Tetlock, 2011). However, we shed light on the drivers behind the reaction to stale news (i.e., stale news is not retelling stories in a neutral way). More specifically, retelling articles with more negative personalization is driving the negative predictive relation with abnormal returns. We control for various article and stock characteristics, including the sentiment of the text (e.g., Tetlock, 2007; Garcia, 2013).

We shift our focus to how transmission bias affects investor's trading behavior. Huberman and Regev (2001) document a sharp increase in trading volume to no-new-news. We show that retelling articles is associated with an increase in abnormal turnover in the following trading month. The positive relation between retelling and abnormal trading volume is driven by more negative personalization and retelling articles by less specialized news outlets. Temporal gaps and competition among news outlets are not driving the relation between retelling articles and abnormal turnover. Additionally, using retail trading data and 13F institutional holdings data, we show that retail traders are driven by retelling articles while institutional traders are not.

An important test in our paper is looking at how distortion from retelling articles impacts the information environment. We examine how retelling articles relate to future analyst forecast dispersion, disagreement among social media users, and the stock's implied volatility. We find a strong positive relation between retelling articles and analyst forecast dispersion in the following trading quarter. Similarly, we show that retelling articles is also associated with increased disagreement among social media users and an increase in the average implied volatility of the underlying stock. This evidence suggests that distortion from retelling articles impacts the information environment of the firm and leads sophisticated (i.e., financial analysts) and unsophisticated (social media users) market participants to disagree. However, if retelling articles are only a way to increase investor attention (e.g., Engelberg and Parsons, 2011; Hillert, Jacobs, and Müller, 2014) or sentiment (e.g., Tetlock, 2007; Hribar and McInnis, 2012; Garcia, 2013; Jiang, Lee, Martin, and Zhou, 2019), one would not expect to see an increase in disagreement, especially among sophisticated participants.

We continue to find a negative link between stocks mentioned in retelling articles and subsequent performance when we perform portfolio sort analysis. We compare the future performance of portfolios sorted on whether a stock is mentioned in a retelling article and consider differences in factor loadings across portfolios. We find that the retelling portfolio generates 1.233% lower monthly alpha compared to the portfolio of stocks not mentioned in retelling articles.

This paper makes three important contributions to the literature. First, we are the first to document transmission bias in financial news by carefully examining how financial news stories can be distorted as they spread in the market by different media outlets. Prior studies have documented various biases in financial media, including a tendency to favor coverage of firms that advertise with them, local firms, and firms that share similar political views (e.g., Reuter and Zitzewitz, 2006; Gurun and Butler, 2012; Goldman, Gupta, and Israelsen, 2024). We argue that bias in the media can originate in the process of disseminating the news itself. We examine how the distortion documented above impacts asset prices, trading behaviors, and the information environment. Prior studies argue that investors act on stale information due to behavioral biases like overconfidence (Tetlock, 2011). However, we show that retelling articles are not simply restating facts from the original source. There are clear patterns to how this information is sensationalized before it gets to the investor. This sensationalization has important financial market implications. In addition, we document some of the mechanisms behind the distortion, including the specialization of the news outlet, the temporal gap between the original story and the retelling article, and competition to retell the story among news outlets. Barber, Huang, Odean, and Schwarz (2022) attribute part of the attention-induced trading behavior of Robinhood traders to the information environment in the Robinhood app. We argue that part of the reaction to stale news stems from the distorted content in retelling articles.

Second, we overcome a limitation in the literature of capturing distortion in news by utilizing ChatGPT-4. Related papers often use methods that do not consider context to measure specific aspects of distortion. For example, prior work used the Loughran and McDonald (2011) dictionary to measure media slant based on tone (Gurun and Butler, 2012), the Jaccard (1901) index to capture text similarity (Tetlock, 2011), and counting the presence of political phrases to capture political slant (Gentzkow and Shapiro, 2010). In contrast, we use ChatGPT-4 to quantify how retelling articles differ from the original story in multiple dimensions. Unlike other machine learning approaches, ChatGPT-4 has a deeper understanding of language context, enabling it to grasp the nuances of sentences and paragraphs. This makes ChatGPT-4 suitable for tasks that require an understanding of whole texts, such as summarizing news articles or answering questions based on them, which are the main tasks of this study. More recently, researchers apply ChatGPT for parsing news headlines (Lopez-Lira and Tang, 2023), Federal Reserve announcements (Hansen and Kazinnik, 2023), and forecasting innovation (Yang, 2023). Our findings demonstrate that ChatGPT-4 can be a useful tool to help us understand how news is distorted as it spreads in financial markets.

Third, we specifically examine the distortion between two related articles, where the second article is a sequential retelling or based on the content of the original article. This helps us specifically capture how news outlets might vary in how they report news and rule out new stories and events that might take place contemporaneously. Gurun and Butler (2012) combine news stories for a given firm from a certain news outlet monthly and compute the proportion of negative words used in the text using a dictionary approach. They find that local media report news more positively about local companies. Our study provides an important contribution to this area in the following ways. First, we ensure that we examine distortion among articles strictly about the same story. This feature makes our setting resemble the "telephone game" where a message changes as it is passed from one party to another. On the other hand, Gurun and Butler (2012) focus is more general on the coverage of firms in various news media. Second, we go beyond measuring differences in tone to capture distortion. Our measure of distortion utilizes ChatGPT-4, which considers differences in tone, factual details, opinions, and appeal of the writing style.

Broadly, our paper is related to research focusing on news and financial markets (e.g., Tetlock, 2007; Fang and Peress, 2009; Engelberg and Parsons, 2011; Obaid and Pukthuanthong, 2022). More specifically, our paper is related to finance and economics studies examining bias or slant in news media (e.g., Gentzkow and Shapiro, 2006, 2008, 2010; Gurun and Butler, 2012). Gentzkow and Shapiro (2010) create an index for media political slant at the news outlet level by comparing the language used in news outlets and *Congressional Record*. They show that media outlets respond strongly to the political preferences of the readers.

Our paper is related to Tetlock (2011). The author calculates the text similarity using the Jaccard (1901) index between news text and the previous ten news stories related to the same firm and argues that this measures stale news and shows that individual investors overreact to stale news. Even though the approach is straightforward and does not require training data, one major limitation of this approach is that the Jaccard (1901) index does not take context and semantic meaning into consideration when computing similarity. You can have two articles with high Jaccard scores that contain new information. Compared to Tetlock (2011), we carefully verify retelling articles to ensure they are retelling the same story as the original article. Unlike Tetlock (2011), our focus is examining how the retelling stories distort the message in the original story (i.e., transmission bias).

Our study is also related to studies focusing on the dissemination of information. For example, Cookson, Engelberg, and Mullins (2023) document patterns consistent with echo chambers among users of a popular finance social media platform. The concept of an echo chamber refers to a situation where individuals are exposed only to information that reinforces their existing beliefs, leading to a narrow perspective. While echo chambers and transmission bias both involve the dissemination of information, our study examines the transformation of content as it passes through various channels, whereas the echo chamber concept centers on the selective exposure and reinforcement of ideas within closed systems.

2. Data

In this section, we discuss the various data we use in this study. First, we discuss the sample of exclusive WSJ articles. Next, we explain the sample of retelling articles from Factiva and the method of identifying which retelling articles are related to the exclusive WSJ articles. We also discuss key variables, including those created based on ChatGPT-4.

2.1. Exclusive WSJ sample

Our study utilizes the full text of all the exclusive WSJ articles. These stories have been tagged "exclusive" by the journal to indicate that the WSJ has information, an interview, or a story that no other media outlet has at the time of publication. This sample ensures that, up to the point of publication, the public was not aware of the story in the exclusive WSJ article. We use the "PMDM" tag to isolate the WSJ articles that are labeled exclusive by Dow Jones. We collect all 28,875 exclusive WSJ articles that can potentially be retold by other news outlets.

2.2. Retelling articles sample

We compile a sample of articles that can be considered candidates for retelling exclusive WSJ articles from Factiva. We narrow our corpus to include only articles published by news outlets listed under "Newspapers: Top US newspapers" to avoid articles published by minor blogs or lesser-known news outlets. Moreover, we ensure to exclude any article published in "The Wall Street Journal" itself or any of the publications by Dow Jones-like "Barron's" and "MarketWatch". To ensure consistency and relevance, we stipulate that all articles must be in English and contain explicit references to the WSJ within the first paragraph of the full text.[9] After applying the filters above, we collect a sample of 3,113 articles between 2013 and 2022 across all the major newspapers in Factiva that can possibly match with exclusive

[9]For a comprehensive enumeration of the search phrases considered, refer to Appendix B

WSJ articles from our sample.

2.3. Matching method

We perform document similarity matching between the exclusive WSJ articles and the Factiva sample of retelling articles using TF-IDF and cosine similarity. TF-IDF allows for the weighting of terms based on their importance, reducing the impact of commonly used words that might not be meaningful for similarity analysis. We combine the text of all articles from both groups for TF-IDF vectorization. This process transforms the articles into vectors, allowing the calculation of cosine similarities between each article in the exclusive WSJ articles group and the Factiva retellings group. Term Frequency (TF) measures how frequently a term appears in a document. It is calculated as:

$$TF(t,d) = \frac{\text{“Number of times term } t \text{ appears in document } d\text{”}}{\text{“Total number of terms in document } d\text{”}} \tag{1}$$

Inverse Document Frequency (IDF) assesses the importance of a term t across all documents, formulated as:

$$IDF(t,D) = \log\left(\frac{N}{|\{d \in D : t \in d\}|}\right) \tag{2}$$

where N is the total number of documents in the corpus D, and $|\{d \in D : t \in d\}|$ is the count of documents containing the term t. The TF-IDF score combines these two measurements:

$$TFIDF(t,d,D) = TF(t,d) \times IDF(t,D) \tag{3}$$

After transforming text into TF-IDF vectors, we calculate the cosine similarity between vectors from the exclusive WSJ articles group and all vectors from the Factiva retellings group. Cosine similarity is defined as:

$$\text{cosine\_similarity}(\mathbf{E},\mathbf{R}) = \frac{\mathbf{E}\cdot\mathbf{R}}{\|\mathbf{E}\|\cdot\|\mathbf{R}\|} \tag{4}$$

where E and R are the TF-IDF vectors of the two documents, $(E \cdot R)$ is the dot product of the vectors, and $||E||$ and $||R||$ are the norms of the vectors E and R, respectively.

For each article in the exclusive WSJ group, we identify the top five most similar articles from the Factiva retellings group (with replacement) within 14 days from the publication date of the exclusive WSJ article. We sort these articles by cosine similarity score. The potential matches are verified manually. We review the content of each WSJ article to grasp its primary focus and then evaluate the relatedness of each Factiva retelling article sequentially. A binary system is employed for assessment: 1 if a Factiva retellings group article directly referenced the content of the WSJ article in a meaningful way, such as mentioning a specific event or detail initially reported by the WSJ, and 0 for unrelated articles.

2.4. OpenAI ChatGPT-4 variables

ChatGPT, developed by Open AI, is a large language model that can take sophisticated tasks and provide detailed and clear answers at a level similar to human experts. The model is based on the Generative Pre-trained Transformer (GPT) series of large language models (LLM). The GPT framework utilizes transformer architectures. Transformers are adept at processing sequential data, notably text. More specifically, transformers are characterized by their ability to discern intricate word relationships within sentences. Pioneering this approach, Google's BERT (Bidirectional Encoder Representations from Transformers) emerged in 2018 as a foundational transformer-based model, garnering significant recognition (Devlin, Chang, Lee, and Toutanova, 2018). Subsequently, OpenAI's release of GPT-3 in June 2020, with its unprecedented 175 billion parameters, and trained on 45TB of data. In 2022, Open AI released ChatGPT-3 with its proficient generation of coherent and comprehensive responses across many knowledge areas. Continuing the evolution of transformer-based architectures, OpenAI introduced GPT-4, an even more advanced iteration of the GPT series. GPT-4 is distinguished by its larger parameter count, enhanced training dataset, and improved fine-tuning techniques, resulting in superior understanding and generation of text

(Wiggers, 2023).[10]

The deployment of LLMs to measure distortion among related articles offers a significant methodological advantage over approaches commonly used in the literature, like the "bag of words" (BoW) (Chen et al., 2023). Attempting to measure distortion by looking at the difference between the number of negative or positive words in two related texts is problematic for multiple reasons. First, text can be distorted in many ways beyond just tone. For example, distortion can involve removing factual details in the retelling article or inserting opinions on top of factual story details. Second, BoW is criticized for being overly simplistic, as it only utilizes term frequency and neglects the richness of information conveyed through word order and contextual relationships. With its billions of parameters, LLMs enable researchers to extract nuanced information from textual data into lower-dimensional information-rich vectors.

We analyze the language style of exclusive WSJ news and matched retelling articles using ChatGPT-4. We focus on the following five writing style dimensions: Fact, Opinion, Negativity, Positivity, and Appeal. Based on Melumad, Meyer, and Kim (2021), we create a prompt for ChatGPT-4 comprising 14 questions. In each of the 14 questions, we use a scale of 1 = "not at all", and 7 = "very much so". The prompt is provided in Appendix C. We follow Open AI's prompt best practices guidelines when designing the prompt.[11] We utilize OpenAI's API to request ChatGPT-4 (version GPT-4-0613) to evaluate the articles in our sample. We send the requests between December 10, 2023, and December 12, 2023. Next, we use the responses for each pair of articles to the 14 questions from ChatGPT-4 to score the text on the five dimensions.

The first variable captures the degree of specific factual details in the text. The first two questions in the prompt capture the degree of specific factual details (7 = "very much so",

[10] Despite its performance, ChatGPT has some limitations when used in academic research. First, its responses can change over time as it is updated with new data and algorithms. However, this limitation can be mitigated as we are transparent about the version and date of the model we use. Second, the answer also depends on the prompt. We follow best practices when constructing the prompt and include it in Appendix C.

[11] Source: Open AI Prompt Engineering

and 1 = "not at all") and the degree of vagueness (reverse-coded, 7 = "not at all", and 1 = "very much so") in the text, respectively. The sum of the ratings on the first two questions is our *Fact* variable. We construct $Fact_A$ and $Fact_{Match}$ that correspond to the exclusive WSJ articles and the matched retelling articles (range between 2 and 14), respectively.

The second variable captures the presence of opinions in the text. For this index, we sum ratings on questions 8, 11 (reverse-coded), and 14 (reverse-coded) as our *Opinion* variable. We construct $Opinion_A$ and $Opinion_{Match}$ that correspond to the exclusive WSJ articles and the matched retelling articles (range between 3 and 21), respectively.

The third and fourth variables capture negativity and positivity in the text. Negativity (positivity) is the degree of opposition (support) or disagreement (agreement) conveyed in the text. *Neg* is the sum of questions 6, 10, and 13. *Pos* is the sum of elements 5, 9, and 12. We construct Neg_A (Pos_A) and Neg_{Match} (Pos_{Match}) that correspond to the exclusive WSJ articles and the matched retelling articles (range between 3 and 21), respectively.

The final variable is capturing the appeal of the text. In other words, how interesting the text is to read and the quality of the writing. We create *Appeal* as the sum of questions 4, 5, and 7. We construct $Appeal_A$ and $Appeal_{Match}$ that correspond to the exclusive WSJ articles and the matched retelling articles (range between 3 and 21), respectively.

Earlier in the paper, we mention two examples of exclusive WSJ articles in our sample that were retold by NYP and WP. The first one is about how Mudrick Capital Management LP experienced a 10% loss in its flagship fund, published on June 11, 2021. The second one is about how PepsiCo Inc. is implementing layoffs at the headquarters of its North American snacks and beverages divisions, published on December 5, 2022. To help understand the reasoning behind the variables above, we ask ChatGPT-4 to identify specific phrases or sections in the articles that justify the ratings assigned to the exclusive WSJ article and the respective retelling.

We provide the detailed response from ChatGPT-4 for the two articles in Appendix D. Overall, ChatGPT demonstrates a robust capability in understanding and evaluating textual content against a detailed rubric, as exemplified in its analysis of two example articles. In both the first original article (Mudrick) and its respective retelling article in the NYP, ChatGPT highlights the extensive use of specific financial strategies and outcomes to justify high scores in detail and the quality of writing while also noting the neutrality or mild critical tones in evaluating the articles' positivity or negativity towards the subject matter. Furthermore, ChatGPT adeptly handles more nuanced criteria, such as the presence or absence of emotional language and the article's stance towards public opinions, showcasing its ability to perform detailed and context-aware evaluations.

The model effectively distinguished between the detailed factual reporting of the second original article (PepsiCo) and the broader economic analysis of the retelling article in the WP. Notably, ChatGPT effectively cited specific text passages to justify each rating, illustrating its ability to ground its evaluations in direct textual evidence.

2.5. Automated text variables

We compute various automated text variables to help us validate our responses from ChatGPT-4 and control for article characteristics in our analysis. To assess the level of complexity and specificity in the text, we compute the following variables. First, we compute Flesch Reading Ease (*readability*), which is a readability test designed to assess the clarity of English writing (Kincaid, Fishburne Jr, Rogers, and Chissom, 1975). [12] A higher score indicates easier readability, with scores typically ranging between 0 to 100. Second, we count the numbers in the text (*numbers*) to assess the level of detail and specificity in the text. Third, we calculate the percentage of words in the text that are complex (*complex*). We define complex words as those with two more syllables. To capture negativity/positivity, we use the 2018 version of the Loughran and McDonald (2011) lexicon. We count the number

[12] The Flesch Reading Ease score is a readability test designed to indicate how easy or difficult a text is to read. The formula for calculating the Flesch Reading Ease score is: Flesch Reading Ease = $206.835 - (1.015 \times \text{ASL}) - (84.6 \times \text{ASW})$, where ASL (Average Sentence Length) is the number of words divided by the number of sentences. ASW (Average Syllables per Word) is the number of syllables divided by the number of words.

of positive words and negative words for each article. We then scale these numbers by the total number of words in the document to get the percentage of positive words and the percentage of negative words (positive, negative). We focus on the net of the negative and positive variables ($neg - pos_{lm}$).

2.6. CRSP variables

We collect returns and volume data from CRSP for companies mentioned by the exclusive WSJ articles and respective retelling articles. We are able to identify 1,330 unique stocks in our sample across 17,984 article-stock observations with return and price data available in CRSP. We calculate the market capitalization (mve_m) of the stock by multiplying the price per share by the shares outstanding. We calculate the market beta ($beta$) for the stock by estimating the Fama and MacBeth (1973) regressions using weekly returns and equal-weighted market returns for three years. Idiosyncratic volatility ($idiovol$) is calculated following Ali, Hwang, and Trombley (2003) by taking the standard deviation of residuals of weekly returns on weekly equal-weighted market returns for a three-year window. We calculate the Amihud (2002) illiquidity measure (ill) by taking the average of daily absolute returns divided by dollar volume. We measure the leverage ratio (lev) following Bhandari (1988) by taking total liabilities and dividing by the market capitalization. We measure the profitability of companies using the definition in Brown and Rowe (2007) by taking the annual earnings before interest and taxes net of non-operating income and scale by non-cash enterprise value ($roic$). Finally, we capture market expectations by calculating the book-to-market ratio following Rosenberg, Reid, and Lanstein (1985) as the book value of equity divided by market capitalization (bm).

We calculate the compounded abnormal returns ($abret$) for the referenced stocks over some window (e.g., $t - 19$ and t or $t + 1$ and $t + 21$) using the Fama-French plus Momentum model (Fama and French, 1993; Carhart, 1997) estimated using a 100-day trading window with a 50-day gap (70-day minimum window). Motivated by Cookson et al. (2023),

$AbLogTurnover$ is the difference between the average daily log turnover between $t+1$ and $t+21$ and the average daily log turnover between $t-140$ and $t-20$ (6-month period, skipping the most recent month).

2.7. Institutional and retail trading data

We use retail trading data based on the method in Boehmer, Jones, Zhang, and Zhang (2021) using TAQ data. The measure takes advantage of a unique aspect of U.S. financial markets regulation: only retail orders, but not institutional orders, can receive price improvement relative to the National Best Bid or Offer (NBBO). We identify orders executed through internalization in TAQ "consolidated tape" as retail buy (sell) if the price is slightly below (above) the round penny. One limitation of this approach is that we can only identify marketable retail orders, but not limit orders, since Reg NMS requires limit orders to be traded at round pennies (Boehmer et al., 2021). Once we calculate the shares of retail orders on a given day, we scale by the shares outstanding on a particular stock to get retail turnover.

We use 13F Holdings Data to calculate institutional trading in stocks in our sample. We aggregate all the shares held by institutions by firm and date across all the institutions. Next, we calculate the quarterly change in shares. Next, we scale by the shares outstanding on a particular stock to get institutional turnover.

2.8. I/B/E/S

We use the Institutional Brokers Estimate System (I/B/E/S) database to obtain the analysts' earnings estimates. The dispersion of analyst forecasts is computed as the standard deviation of EPS forecasts divided by the absolute value of the mean EPS forecast. We use the last estimate for a firm-date observation within a broker-analyst group.

2.9. OptionMetrics

We use the OptionMetrics database to gather data on U.S. options and their implied volatility. We calculate average implied volatility based on call and put options for each stock in our sample using near the money options. We use options between 0.95 and 1.05 moneyness with a minimum of 7 days till the expiration date and a maximum of 186 days till the expiration date. We calculate implied volatility separately for call and put options. We average implied volatility across option contracts and weight by open interest.

3. Results

In this section, we provide our main empirical results. First, we report descriptive statistics for the key variables. Second, we report evidence of transmission bias. More specifically, we provide evidence consistent with disagreeable personalization as exclusive WSJ articles are spread and retold in various news outlets. Third, we examine the mechanism behind the disagreeable personalization. Fourth, we examine how transmission bias can have asset pricing implications. Finally, we examine how transmission bias affects trading behavior and the information environment.

3.1. Descriptive statistics

Table 1 displays descriptive data on exclusive news articles from the WSJ and the subsequent retelling articles by different news organizations. The dataset contains 1,940 observations of exclusive WSJ articles matched to a retelling article.[13] There are a total of 1,351 unique exclusive WSJ articles that are retold. The dataset also contains 27,524 exclusive WSJ articles that are not retold.[14] The sample spans between 2013 and 2022, with a peak

[13]An original story could be retold multiple times.

[14]Our aim in constructing the dataset is to accurately track the retelling of exclusive WSJ articles. By focusing on retelling articles in only the major US newspapers, ensuring retelling articles explicitly reference the WSJ, and carefully confirming the matches, we ensure that our dataset is accurately capturing retelling articles of the same exclusive WSJ story. If we relax the criteria, we increase our sample of exclusive WSJ

in 2014 and the fewest observations in 2017. Out of the 1,351 unique exclusive WSJ articles that are retold, 332 of those articles mention specific tickers that can be matched to CRSP and Compustat variables. In contrast, out of the 27,524 exclusive WSJ articles that are not retold, 9,131 of those articles mention tickers that match CRSP and Compustat variables.

The retelling articles are sourced from 18 different news organizations. The organizations with the most retelling articles are 'INVD (*Investor's Business Daily*)' with 589 retelling articles, followed by 'NYPO (New York Post)' with 356, 'WPCO (*Washington Post*)' with 226, 'USAT (*USA Today*)' with 212, and 'NYDN (*New York Daily News*)' with 98.[15]

The number of companies featured in each exclusive WSJ article that is retold is 2.654 on average, with a median of 2 and a standard deviation of 3.572. For exclusive WSJ articles that are not retold, the average number of tickers mentioned is 1.596 articles, with a median of 1 and a standard deviation of 1.147. The average number of retelling articles per exclusive WSJ article is 1.404× with a standard deviation of 0.818. The average number of days between the publication of an exclusive WSJ article and a retelling article is 2.635, with a median of 1 and a standard deviation of 3.38.

Table 2, Panel A, displays descriptive statistics for the ChatGPT-4 indices of language style. Panel B reports the mean difference between the exclusive WSJ articles and the retelling articles. We note that exclusive WSJ articles are deemed to contain more factual details ($Fact$) and are less vague than the matched retelling articles. More specifically, retelling articles score 12.23% lower, on average, than the original in $Fact$. Retelling articles are judged to be more opinionated ($Opinion$) than the exclusive WSJ articles. Moreover, retelling articles tend to have more negativity (Neg) and less positivity (Pos) than the exclusive WSJ articles. Finally, the exclusive WSJ articles have stronger appeal ($Appeal$) than the retelling articles. All the differences are statistically significant at the 1% level except for Pos, which is significant at the 10% level. The findings in Table 2 support the

articles with a potential matched retelling article; however, the match might discuss the same topic as the exclusive WSJ article, but not exactly the same story and thus not a retelling.

[15] Please see Appendix A for a complete list.

hypothesis that retelling articles goes beyond relaying the key facts in the original story. Retelling articles attempt to provide guidance persuasively by becoming more opinionated and negative and less factual and appealing (as the text becomes more disagreeable). When we calculate the average negative personalization scores ($Opinion - Fact + Neg - Pos - Appeal$) for retelling and original articles and calculate the difference, we find that retelling articles receive a negative personalization score that is, on average, 6.99% higher than the original articles.

In Table 3, we look at the descriptive statistics for the automated text variables that help capture some of the writing styles of the articles. We compute the Flesch Reading Ease of the text to capture linguistic complexity ($readability$) and find that the average readability for the exclusive WSJ articles and retelling samples are 57.225 and 51.108, respectively. Moreover, we look at the presence of numbers ($numbers$) and complex ($complex$) words in the text. Exclusive WSJ articles use an average of 14.814 numbers in the text, whereas the retelling articles use an average of 9.880 numbers in the text. Exclusive WSJ articles contain 0.17% of complex words in the text, whereas retelling articles contain 0.162% of complex words. Text in both samples tends to have more negative than positive words ($neg - pos$), but the retelling articles sample has a slightly more negative tilt than the exclusive sample. The average $cosinesimilarity$ between the articles in both samples is 0.704.

Next, we report the descriptive statistics for the characteristics of the firms mentioned in the exclusive WSJ articles in Table 4, Panel A. We have 16,863 stock-exclusive WSJ article or -retelling article observations in our sample. 6.1% of those article-stock observations are retold by other newspapers. 7.1% represent the retelling articles by newspapers other than the WSJ. We note that the average market beta ($beta$) of stocks in our sample is 1.118, and idiosyncratic volatility ($idiovol$) is 3.5%. The average market capitalization (mve_m) of companies in our sample is \$139 billion. Firms in our sample have an average return on invested capital ($roic$) of 9.5% and a book-to-market ratio (bm) of 0.481. The average leverage ratio (lev) is 2.807%. The average illiquidity using the Amihud (2002) measure is

0.002.

In Table 4, Panel B, we report the descriptive statistics for the stock returns and turnover of firms mentioned in our sample of exclusive WSJ articles. The returns (ret) and abnormal returns ($abret$) corresponding to the month after ($t + 1$, $t + 21$) the publication of the exclusive WSJ article or retelling article is 1.12% and –0.23 bps, respectively. We also report the average abnormal log turnover of stocks in our sample ($AbLogTurnover$) during the time of the article and the month after.

In Table 4, Panels C and D, we compare the firm characteristics and market variables between the stocks mentioned in exclusive WSJ articles that were retold and those that were not. We note that firms in exclusive WSJ articles that are not retold tend to be riskier, less liquid, more levered, have lower profitability, have lower market expectations, and are smaller than those mentioned in exclusive WSJ articles that are retold. Firms mentioned in exclusive WSJ articles that are not retold have lower returns between $t-1$ and t, but higher returns between $t+1$ and $t+21$ compared to firms mentioned in exclusive WSJ articles that are retold.

In Appendix F, we examine the determinants of retelling articles. We show that retelling articles tend to reference more tickers than non-retelling articles. On the other hand, retelling articles tend to be more difficult to read, contain fewer numbers, and fewer complex words. Regarding the stocks that these retelling articles reference compared to non-retelling articles, the companies tend to be less risky, have higher market expectations, and have higher recent abnormal returns.

3.2. Disagreeable personalization

We quantify transmission bias by examining whether information in exclusive WSJ articles is systematically altered in a strategic manner as it is retold across news outlets. On the one hand, news outlets who retell exclusive WSJ articles might simply retell what they deem is an important story that is beyond their purview to their audience. In this case, one would not expect the writing in the retelling article to be noticeably different from the original story, on average. On the other hand, news outlets might attempt to guide and persuade readers in a certain direction. In this case, the retelling article will contain more opinions and guidance. Moreover, the news outlet might emphasize the negative aspects of the content when retelling news, driven by its desire to stand out and persuade the audience to value its guidance (Stapert and Clore, 1969). In the latter case, the final narrative received by the end audience is noticeably different from the original news, on average, characterized by fewer factual details, more opinions, and a more pessimistic tone.

Several factors can influence how information passing across various news outlets is altered. First, memory can influence how stories are recounted, with research suggesting that unusual or unexpected details are more easily remembered and retold (Barrett and Nyhof, 2001). Second, bias can be generated from social and motivational factors (Hyman, 1994; Barasch and Berger, 2014; Melumad et al., 2021). More specifically related to news outlets, advertising pressure (Reuter and Zitzewitz, 2006), media ownership (Besley and Prat, 2006), and competition (Baron, 2005; Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006) can all influence the bias of news outlets. Finally, the specialization of news outlets can shape their retelling of financial stories. For example, more specialized news outlets tend to provide in-depth analysis, while general news outlets simplify financial information for broader appeal.

To examine how retelling articles might change details and language style, we regress the ChatGPT-4 variables of language style on an indicator variable for whether the story is an exclusive WSJ story or a retelling article (*retelling*). Each observation in our regression is an article (either an exclusive WSJ article or a retelling article). We create two additional variables based on the ChatGPT-4 language variables. First, $Opinion - Fact$, which is the difference between $Opinion$ and $Fact$ variables described in Section 2.4 and captures the level of opinions less level of factual details in the text. Second, $Neg - Pos$, which is the difference between Neg and Pos, captures the net pessimism in the text. Finally, we use the logarithm of *Appeal* to capture how appealing the text is perceived.

We also include controls for article characteristics in our regression. First, we control for the number of tickers mentioned in the article ($log(TickerCount)$). It is possible that articles mentioning many tickers are discussing an industry broadly and thus are less focused and detailed, whereas articles mentioning few tickers are focusing on a specific company and thus are more specialized and detailed. We also control for readability using the Flesch Reading Ease (*readability*), the presence of numbers in the text (*numbers*) to assess the level of details and specificity in the text, the percentage of words in the text that are complex (*complex*), and the net % of negative words minus positive words according to the 2018 version of the Loughran and McDonald (2011) lexicon ($neg - pos_{lm}$). We also control for *retold* which is an indicator variable to differentiate between WSJ exclusive articles that are retold and those that are not retold by other news outlets. We run the following regression:

$$LanguageIndex_{i,t} = \alpha + \beta_1 \times retelling_{i,t} + \beta_2 \times STORY_{i,t} \tag{5}$$

where *LanguageIndex* is one of the three ChatGPT-4 language style variables for article i at time t: $Opinion - Fact$, $Neg - Pos$, and $log(Appeal)$. *STORY* is a vector of story characteristics including number of tickers mentioned, readability score, % of negative net positive words in the text, count of numbers in the text, and % of complex words. We include time-fixed effects and cluster standard errors by article ID. We report the results in Table 5.

In the first specification, we find a strong and significant (at the 1% level) positive relation between *retelling* and $Opinion - Fact$, suggesting that the retelling article tends to, on average, contain more opinions and less factual details than the original story. This relation is consistent with alterations or exaggerations being introduced as news outlets add their interpretations or perspectives to the original story. In the second specification, we include story-level controls and continue to find a positive and significant relation between *retelling* and $Opinion-Fact$. Moreover, we find a negative and significant coefficient (at the 1% level) on $numbers$ suggesting that text deemed to be more opinionated and less factual by ChatGPT contains fewer numbers. This relation is intuitive, as one would expect more opinionated writing to contain fewer numbers. Hence, the negative coefficient on $numbers$ helps validate our $Opinion - Fact$ from ChatGPT-4.

In the third specification, we find a positive and significant (at the 1% level) relation between *retelling* and $Neg-Pos$, indicating that retelling articles tend to be more pessimistic than the original. In the fourth specification, we include story-level characteristics as controls and continue to find a positive and significant (at the 1% level) relation between *retelling* and $Neg - Pos$. Consistent with Stapert and Clore (1969), news outlets are leveraging negativity to capture the attention of readers. The coefficient on $neg - pos_{lm}$ is positive and highly significant, suggesting that the automated pessimism measure from the 2018 version of the Loughran and McDonald (2011) lexicon and the ChatGPT-4 index are related in the predicted way. This relation helps validate our $Neg - Pos$ ChatGPT-4 variable.

In the fifth specification, we focus on the natural logarithm of *Appeal* (*log*(*Appeal*)), which captures how interesting the text is to read and the quality of the writing. We find a negative and significant (at the 1% level) coefficient on *retelling* suggesting that retelling articles are deemed to be less interesting to read with lower writing quality compared to the original exclusive WSJ articles. In the sixth specification, we add story-level characteristics. We continue to find a negative and significant coefficient on *retelling*. Moreover, we show a negative and significant (at the 5% level) relation between *readability* and *log*(*Appeal*) which suggests that text that is more challenging to read is less interesting to read with lower writing quality, on average. One would expect that the readability score (a lower score suggests the text is more difficult to read) will be lower for text that is less appealing to read. Hence, the negative relation between *readability* and *log*(*Appeal*) helps validate our *Appeal* ChatGPT-4 measure. After we include all the article-level characteristics and time fixed effects, we see R^2 between 4.9% and 21.1% suggesting that ChatGPT can capture information about how two articles about the same story are different beyond what the automated text variables can capture.

Overall, in Table 5, we provide evidence that retelling articles tend to be more disagreeable, opinionated, and less interesting to readers than the original. The results are statistically strong and hold after controlling for various article-level controls. Our evidence suggests that memory is not the only factor present in how news is retold. Our evidence suggests that news is retold systematically in a strategic way. Moreover, we provide evidence to help validate our ChatGPT-4 language variables.

3.3. Mechanisms: Specialization, memory, and motivation

In this section, we aim to explore the underlying mechanisms behind the systematic shift in news content as it is retold. We explore how specialization, memory, and social and motivational factors can play a role in how the content might shift as it is retold.

First, we compare how news outlets that rarely versus those that frequently retell stories by the WSJ shift the content of the original story. We postulate that news outlets that rarely retell news from other organizations are likely more specialized and thus rely on their own reporting for the most part. In contrast, news outlets that frequently retell news from the WSJ are possibly not specialized and regularly depend on outside sources to cover financial topics. When news outlets perceive themselves as more knowledgeable and specialized about a specific topic, they are more inclined to include their own opinions and interpretations.

Second, we explore how the time between the original story and the time of the retelling impacts how the story shifts compared to the original. A longer temporal gap creates more room for distortion as the story starts to fade from the memory of the public (Barrett and Nyhof, 2001). This gap allows for the accumulation of errors, misinterpretations, and the insertion of bias as the story is passed from one journalist to another.

Third, we explore how distortion varies between exclusive WSJ articles that are retold by many news outlets versus those that are only retold by a few. Competition between news outlets can bias their reporting (Baron, 2005; Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006). When an exclusive article from WSJ is retold by several other news outlets simultaneously, the potential for distortion increases due to the tendency of reporters to sensationalize the story to gain more readers in a competitive environment.

In Table 6, we regress the ChatGPT-4 language style variables on three key variables behind retelling articles. First, to compare distortion by news outlets that rarely or frequently retell news from the WSJ, we $uselog(NewsOrgFreq)$, the natural logarithm of the number of retelling articles by the news outlet in the past. Second, to capture how time since the original article can impact distortion, we use $log(DaysBetween)$, which is the natural logarithm of the days between the date of the original exclusive WSJ article and the retelling article. Third, to capture how the number of competing retelling articles of the same story on the same day can impact distortion, we use $log(RetellingsSameDay)$, which is the natural logarithm of the number of competing retelling articles on the same exclusive WSJ article on the same date.

In the first specification of Table 6, the dependent variable is $Opinion - Fact$. We find a positive and significant (at the 1% level) coefficient on $log(NewsOrgFreq)$, indicating that articles retelling exclusive WSJ stories by news outlets that frequently retell WSJ stories tend to be more opinionated and less factual. This finding is consistent with less specialized news outlets (those that retell news often) injecting more opinions and removing factual details in their retelling. Next, we find a positive and significant (at the 1% level) coefficient on $log(DaysBetween)$. This finding suggests that the larger the time-lapse between the original and the retelling articles, the retelling articles tend to be more opinionated and less factual. This finding is consistent with the temporal gap allowing for accumulation of bias. We do not find a significant relation between competition among retelling stories and the level of opinions and factual details in the writing of retelling articles.

In the second specification of Table 6, we focus on the level of disagreeableness in the text. We find a negative and significant (at the 10% level) on $log(NewsOrgFreq)$, which suggests that news outlets that frequently retell exclusive WSJ articles (less specialized) tend to write articles that are less disagreeable. Moreover, we document a significant (at the 1% level) and positive relation between the time-lapse between the retelling article and the original story and the level of disagreeableness in the retelling article. Next, we look at $log(RetellingsSameDay)$ and find a positive and significant (at the 1% level) relation with $Neg - Pos$. This finding is consistent with news outlets relying on negativity to capture the attention of readers in a more competitive environment.

In the third specification of Table 6, the main dependent variable is $log(Appeal)$. The coefficient on $log(NewsOrgFreq)$ is negative and significant at the 1% level. This result is consistent with news outlets that retell often producing articles that are less appealing to read. The coefficient on $log(DaysBetween)$ is not statistically significant. We find a positive and significant coefficient on $log(RetellingsSameDay)$, which is consistent with competition encouraging news outlets that want to retell the exclusive WSJ story to write a more interesting and appealing article.

In Table 6, we document that news outlets that frequently retell exclusive WSJ stories tend to write more opinionated, less disagreeable, and less appealing articles. We document that the larger the time-lapse between the retelling article and the original results in retelling articles that are more opinionated and disagreeable. Finally, when there is a competition to retell exclusive WSJ stories, retelling articles tend to be, on average, more disagreeable in tone and more appealing in the hope of catching attention. Overall, we document important factors that help determine how financial stories are retold.

3.4. Abnormal returns

The retelling of financial news, especially when news is distorted or interpreted differently by various newspapers, can impact asset returns. Fang and Peress (2009) argue that media coverage reaches a broad audience and impacts expected returns. Moreover, Tetlock (2010) shows that media coverage reduces information asymmetries. In this section, we explore how transmission bias can impact future asset returns. Retelling articles with higher negative personalization scores can negatively impact investors' perception of the original story, leading to lower future returns. In contrast, retelling articles with higher negative personalization scores can attract greater attention from investors and the media, increasing the visibility of the stock and disagreement among investors. Disagreement among investors can push prices up, especially in light of short-sale constraints (e.g., Diether, Malloy, and Scherbina, 2002; Jones and Lamont, 2002.

We regress next month's ($t+1$, $t+21$) abnormal returns for stocks mentioned in the exclusive WSJ article or retelling article on indicators for whether the article has been retold (*retold*) or is a retelling of an exclusive WSJ article (*retelling*). Each observation in our regression is an article-firm. Moreover, we add 1) $STORY$, which is a vector of story characteristics including number of tickers mentioned, readability score, % of negative words in the text, count of numbers in the text, and % of complex words; and 2) $FIRM$ which is a vector of observable firm characteristics including industry fixed effects and past returns. We cluster standard errors by firm.

$$log(abret_{i,j,t,t+21}) = \alpha + \beta_1 \times retold_{i,j,t} + \beta_2 \times retelling_{i,j,t} + \beta_3 \times STORY_{i,j,t} + \beta_4 \times FIRM_{i,j,t-1} \tag{6}$$

where $abret_{i,j,t,t+21}$ is the compounded abnormal returns between $t+1$ and $t+21$ adjusted using the Fama-French three-factor model plus momentum for stock j mentioned in article i. We use a 100-day estimation window 50 days prior to time t and require a minimum of 70 valid returns to estimate the expected returns.

We report the regression results in Table 7. In specification (6), instead of a retelling variable, we add more nuanced variables to capture specific characteristics of the retelling article, including 1) NegPer; 2) log(NewsOrgFreq), 3) log(DaysBetween); and 4) log(Retellings SameDay). These variables are 0 for all articles except those that are retelling articles of an exclusive WSJ article. $NegPer$ is the difference between the negative personalization scores between the original story and the retelling. The negative personalization score is calculated using the ChatGPT-4 language style indices as: $Neg - Pos + Opinion - Fact - Appeal$. A higher value of $NegPer$ indicates that a retelling article has more negative personalization compared to the original exclusive WSJ article (more opinionated, more negative tone, less factual, less positive tone, and less appealing). $log(NewsOrgFreq)$ captures the frequency of retelling articles by various news outlets. We postulate that a higher value means that the news outlet is less specialized and thus relies less on their own reporting. To differentiate between retelling articles that occur soon after the initial report and those that come later, we include $log(DaysBetween)$, which is the log of the number of days between the retelling article and the exclusive WSJ article. Finally, we capture whether there are competing retelling articles about the same exclusive WSJ article on the same day using the $log(RetellingsSameDay)$.

Before focusing on the predictability results, we show how *retelling* relates to contemporaneous monthly abnormal returns in specifications (1) and (2) in Table 7. We find no significant relation between retelling articles and contemporaneous monthly abnormal returns. Starting in specification (3), we focus on the predictability results. We find a negative and significant coefficient on *retelling*.[16] Once we control for story and firm characteristics in specifications (4) and (5), we continue to find that retelling articles negatively predict abnormal returns. The relation is statistically significant at the 1% level after controlling for firm and story characteristics.

In specification (6), we focus on the factors behind the retelling articles. Instead of the *retelling* dummy, we focus on 1) $NegPer$; 2) $log(NewsOrgFreq)$, 3) $log(DaysBetween)$; and 4) $log(RetellingsSameDay)$. We find that only $NegPer$ is negative and significant, while the other three factors are insignificant. In other words, *retelling* articles with higher negative personalization are associated with lower future abnormal returns. This result suggests that negative personalization mostly drives the negative predictive relation between

[16]We check for multicollinearity issues in our regression and find an average variance inflation factor (VIF) of 1.79 and none of the VIF is larger than 10.

retelling articles and future abnormal returns.

Overall, our evidence is consistent with negative personalization in retelling articles, which gives the public a gloomier perception of the original story, leading to lower next month's abnormal returns.

3.5. Turnover

Next, we examine how transmission bias can generate trading. As news stories are circulated and altered, they capture the attention of a broader audience and can increase disagreement. Disagreement among investors can lead to an increase in trading (Karpoff, 1986). More recently, Goldman et al. (2024) show that politics-induced disagreement in the coverage of corporate news is associated with increase in abnormal trading volume. In Table 8, we focus on the relationship between retelling and future abnormal share turnover. To examine this relation, we run the following regression:

$$AbLogTurnover_{i,j,t+1,t+21} = \alpha+\beta_1\times retold_{i,j,t}+\beta_2\times retelling_{i,j,t}+\beta_3\times STORY_{i,j,t}+\beta_4\times FIRM_{i,j,t-1} \tag{7}$$

where $AbLogTurnover_{i,j,t+1,t+21}$ is the difference between average daily log turnover between $t+1$ and $t+21$ and the average daily log turnover between $t-140$ and $t-20$. We also control for firm and story characteristics. If retelling articles are responsible for generating trading, we expect to find $\beta_2 > 0$. In specification (4), instead of *retelling* variable, we add: 1) $NegPer$; 2) $log(NewsOrgFreq)$, 3) $log(DaysBetween)$; and 4) $log(RetellingsSameDay)$.

In Table 8, we find a positive and statistically significant (at the 1% level) relation between *retelling* and abnormal share turnover in the following month. This finding is consistent with retelling articles generating more disagreement among traders and leading to an increase in trading. In specification (4), we examine which characteristics of the retelling article are driving the positive relation between *retelling* and future abnormal trading volume. We note a positive and significant (at the 1% level) coefficient on $NegPer$, highlighting that the negative personalization drives much of the abnormal trading volume. Moreover, retelling articles by news outlets that retell more often (less specialized) are associated with higher abnormal trading volume. The less specialized news outlets might have a larger and less sophisticated audience and thus lead to higher turnover.

3.6. Retail and institutional trading

In this section, we explore how retail and institutional traders behave differently in response to retelling news. Consistent with papers portraying retail traders as unsophisticated "noise" trades in the behavioral finance literature (Barber and Odean, 2008; Barber, Odean, and Zhu, 2008; Barber and Odean, 2000; Barber, Lee, Liu, and Odean, 2009), one would expect that retail traders will respond more strongly to transmission bias. Whereas institutions are more sophisticated with better tools and thus are less likely to act on distortions in retelling articles. However, if retelling articles are contributing to a more nuanced and complete information environment, one can expect that institutions will respond to retelling articles.

In Table 9, we look at the relation between *retelling* and future abnormal share turnover for retail and institutional traders. To examine this relation, we run the following regressions:

$$\begin{aligned} AbLogTurnoverRetail_{i,j,t+1,t+63} \text{ or } AbLogTurnoverInst_{i,j,t+1,t+63} &= \alpha + \beta_1 \times retold_{i,j,t} \\ &+ \beta_2 \times retelling_{i,j,t} + \beta_3 \times STORY_{i,j,t} + \beta_4 \times FIRM_{i,j,t-1} \end{aligned} \tag{8}$$

where $AbLogTurnoverRetail_{i,j,t+1,t+63}$ ($AbLogTurnoverInst_{i,j,t+1,t+63}$) is the difference between average daily log turnover for retail (institutional) traders between $t+1$ and $t+63$ and the average daily log turnover for retail (institutional) traders between $t-252$ and $t-63$ (9-month period, skipping most recent quarter).[17] We also control for firm and story

[17] The reason we use quarterly windows instead of monthly or daily is that we are limited by the quarterly frequency in the 13F holdings data.

characteristics.

In specification (1) of Table 9, we focus on abnormal turnover for retail investors and find a positive and significant coefficient on *retelling* (at the 5% level). In specification (2), we focus on institutional trading and find no significant relation between future abnormal turnover and *retelling*. Overall, we find that retail traders are influenced by retelling articles but not institutional trading.

3.7. Information environment

We explore whether distortion from retelling articles is associated with uncertainty in the information environment. Distortion from retelling articles can impact investors' information environment differently and generate disagreement among investors. More specifically, we explore how retelling articles relate to future disagreement among analysts and social media users, and average implied volatility. Bailey, Cao, Kuchler, and Stroebel (2018) and Cookson and Niessner (2020) examine the differences among investors that lead to disagreement and how this disagreement impacts trading behavior. Even though researchers find that disagreement from differences in information environments and trading models both impact trading outcomes; our focus is on the role bias in the media plays in the information environment of traders.

One would expect that if retelling articles add distortion and confusion to the information environment, disagreement among investors will increase. In contrast, if retelling articles are simply a way to increase investor attention (e.g., Engelberg and Parsons, 2011; Hillert et al., 2014) or sentiment (e.g., Tetlock, 2007; Hribar and McInnis, 2012; Garcia, 2013; Jiang et al., 2019), one would expect that disagreement would not be impacted, especially among sophisticated participants. Alternatively, if retelling articles add clarity to the original exclusive WSJ story, one would expect disagreement to drop.

We first test this prediction relating to disagreement among analysts in Table 10. We regress dispersion in analyst forecasts in the quarter after the article is published *Analyst*

$Disper_{t+1,t+63}$ on indicators for whether the article is a retelling of an exclusive WSJ article (*retelling*).

$$log(AnalystDisper_{i,j,t+1,t+63}) = \alpha + \beta_1 \times retold_{i,j,t} + \beta_2 \times retelling_{i,j,t} + \beta_3 \times STORY_{i,j,t} + \beta_4 \times FIRM_{i,j,t-1} \tag{9}$$

where $AnalystDisper_{i,j,t+1,t+63}$ is the standard deviation of analyst forecasts made in the quarter post article publication (between $t+1$ and $t+63$).

In the first specification of Table 10, we note a positive and significant (at the 1% level) coefficient on *retelling*. In other words, retelling articles lead to an increase in analyst disagreement. In the second specification, we add firm-level controls and continue to observe a similar result. Finally, in specification (3), we add article characteristics and our results hold. Our findings in Table 10 are consistent with a distortion story instead of an information or sentiment story.

We shift our focus to disagreement among social media users. We use data on attention and disagreement among users of the social media platform StockTwits Cookson and Niessner (2020). We replace the dependent variable from the regression above with $Dis_{i,j,t+1,t+21}$ in Table 11. $Dis_{i,j,t+1,t+21}$ is the natural logarithm of the standard deviation of the sentiment across social media posts about stock j. In Table 11, we note that retelling articles are associated with an increase in future disagreement among social media users on StockTwits. The relation is statistically and economically strongest for within-group disagreement (informational) compared to cross-group disagreement (investment philosophy).

Next, we run a similar regression to the one in Table 10, except we replace $AnalystDisper_{t+1,t+63}$ with either $ImpliedVolCall_{t+1,t+21}$ and $ImpliedVolPut_{t+1,t+21}$ in Table 12. In the first three specifications, we focus on average future implied volatility using call options and show a positive and significant coefficient on *retelling* (significant at the 1% level after adding firm and article characteristics). We observe identical results when we look at implied volatility based on put options instead of call options.

Overall, our evidence on the information environment is consistent with transmission bias, which creates differences in readers' information environment and leads to an increase in disagreement and attention among investors.

3.8. Portfolio sort analysis

We perform portfolio sort analysis to help provide robustness and demonstrate the economic value of our findings on return predictability. We sort stocks into two portfolios based on whether the stock is mentioned in a retelling article (*Retelling* portfolio) or not (*NoRetelling* portfolio). We rebalance portfolios daily and hold the stocks in the portfolios between $t+1$ and $t+21$.

We take the average daily return in each portfolio and test the difference between the *Retelling* and *NoRetelling* portfolios. We restrict our sample to days with at least 10 stocks in each portfolio. Moreover, the time-series t-statistics are based on Newey-West standard errors (with five lags). we report our results in Table 13. We find that the *Retelling* portfolio has a 0.061% significantly (at the 5% level) lower daily return than the *NoRetelling* portfolio (-1.274% monthly).

Next, we report the performance for the retelling-sorted portfolios after considering Carhart (1997) factor exposures. In the bottom half of Table 13, we summarize the results from the test for the one-month performance window. We find a statistically and economically strong negative relation between *retelling* and future stock performance. The *Retelling* portfolio generates 5.9 basis points lower daily alpha than the *NoRetelling* portfolio (-1.233% monthly).

3.9. Reversal

In this section, we check how long the predictability results we document in Table 7 last. If we note a reversal in the following month, the predictability results are likely due to temporary mispricing in the market from transmission bias. In contrast, if we see a continuation of the predictability results, it suggests that the retelling articles might disclose or highlight risk factors that the initial story did not highlight or mention clearly.

We run the same model as model 5 in Table **??**, except we replace $abret_{i,j,t,t+21}$ with $abret_{i,j,t,t+42}$ and $abret_{i,j,t,t+63}$ to check whether the predictability results disappear after we extend the predictability horizon to 2 and 3 months, respectively. We show that in models 2 and 3 of Table 14, the coefficients on *retelling* are no longer statistically significant. This evidence supports the temporary mispricing story.

3.10. Tone, personalization, and appeal

We breakdown $NegPer$ into its three components and examine which of these components drive our return predictability results in Table 7. We create two new variables: $Tone$ and $Personalization$. $Tone$ is the difference between the tone between the original story and the retelling article. The tone is calculated as $Neg - Pos$. $Personalization$ is the difference between the personalization scores between the original and retelling articles. The personalization score is calculated as $Opinion - Fact$. A higher$Tone$ value indicates that a retelling article has a more negative tone than a positive tone compared to the original exclusive WSJ article. A higher value of $Personalization$ indicates that a retelling article has more opinions than facts than the original article. We regress future abnormal returns on the two new variables above plus *Appeal* as our primary independent variables, plus article and firm characteristics. We report our results in Table 15.

The negative relation between $NegPer$ and future abnormal returns documented earlier is mostly driven by the personalization of the story as opposed to the increase in the negative tone or decrease in the appeal. We show that $Personalization$ has a negative and significant coefficient, while the coefficient on $Tone$ is not statistically significant. In the third specification, we focus on *Appeal* and show a positive and significant coefficient (at the 10% level). Overall, the increase in opinions and decrease in factual details and appeal of the retelling articles relative to the original article is driving the negative relation between $NegPer$ and future abnormal returns.

4. Conclusion

This study advances the literature by studying how financial news is distorted as it proliferates through the market, challenging the prevailing assumption that investor behavior is predominantly influenced by acting on outdated information due to cognitive biases such as overconfidence. We argue that distortion in media as it spreads is partly responsible. First, We demonstrate that the spread of news stories goes beyond mere repetition, revealing systematic patterns of sensationalization prior to reaching investors. Second, we identify several factors driving the distortion of original stories, including the news outlet's specialization, the time-lapse between the initial story and its retelling article, and competitive pressures among media outlets. Third, we examine the implications of these distortions for asset returns, trading behavior, and the overall information landscape. Mainly, we show that transmission bias negatively predicts abnormal monthly returns and positively predicts abnormal trading volume and disagreement among investors. Fourth, we employ the advanced capabilities of ChatGPT-4 to quantify transmission bias across various dimensions.

Appendix A. News outlets

News ID	N	News Outlet
INVD	589	Investor's Business Daily
NYPO	356	New York Post
WPCO	226	Washington Post.com
USAT	212	USA Today
NYDN	98	New York Daily News
BSTN	80	The Boston Globe
WP00	66	The Washington Post
NYTF	65	NYTimes.com Feed
CHSM	62	The Christian Science Monitor
AMB0	52	American Banker
PPGZ	44	Pittsburgh Post-Gazette
SLMO	34	St. Louis Post-Dispatch
STPT	17	Tampa Bay Times
MSP0	15	STAR TRIBUNE (Mpls.-St. Paul)
ATJC	15	The Atlanta Journal
PHLI	6	The Philadelphia Inquirer
BFNW	2	Buffalo News
NYTA	1	The New York Times Abstracts

Appendix B. Search phrases

Search phrases to find retelling articles in Factiva
According to the Wall Street Journal
appeared in The Wall Street Journal
article in The Wall Street Journal
by The Wall Street Journal
information of The Wall Street Journal
quoted in the Wall Street Journal
report from The Wall Street Journal
report in The Wall Street Journal
reported by the Wall Street Journal
reports the Wall Street Journal
Sources tell The Wall Street Journal
Speaking to the Wall Street Journal
story in The Wall Street Journal
The Wall Street Journal first reported
The Wall Street Journal had reported
The Wall Street Journal initially reported
The Wall Street Journal is reporting
The Wall Street Journal looks
The Wall Street Journal previously reported
The Wall Street Journal publishes
The Wall Street Journal say
The Wall Street Journal writes
told the Wall Street Journal
Wall Street Journal claimed
Wall Street Journal claims
Wall Street Journal described
Wall Street Journal describes
Wall Street Journal found
Wall Street Journal has reported
Wall Street Journal published
Wall Street Journal report
Wall Street Journal reported
Wall Street Journal reports
Wall Street Journal revealed
Wall Street Journal said
Wall Street Journal says

Appendix C. ChatGPT-4 prompt

You will be provided two related news articles. These two articles (delimited with XML tags, the first article has the ¡firstarticle¿ tag, and the second one has the ¡secondarticle¿) come from different newspapers but are about the same story. Your task is to read and rate the two news articles. First, read the first article and rate it (on a 1: "not at all" to 7: "very much so" scale) based on how much you agree with the following statements related to the text:

1. This text is detailed;
2. This text is general / vague;
3. This text is interesting;
4. This text is well-written;
5. This text is positive about the subject matter;
6. This text is negative about the subject matter;
7. This text conveys interest in the subject matter;
8. This text is emotional about the subject matter;
9. In this text, the writer expressed support for the subject matter at hand;
10. In this text, the writer took an opposing stance towards the subject matter at hand;
11. In this text, the writer expressed no opinion about the subject matter itself;
12. In this text, the writer agreed with other people's opinions about the subject;
13. In this text, the writer disagreed with other people's opinions about the subject;
14. In this text, the writer expressed no opinion about other people's opinions.

Next, read the second article and rate it like above. However, this time when rating the second text, please think about how it compares to the first text (e.g., relative to the first article, the second article is more... or relative to the first article, in the second article the writer...:).

Your response should be in this format [firstarticle = X, X, X, X, X, X, X, X, X, X, X, X, X, X, secondarticle = X, X, X, X, X, X, X, X, X, X, X, X, X, X], where X is a number between 1 and 7 that corresponds to the rating on the 14 elements above.

Appendix D. ChatGPT-4 example responses with justification

We ask ChatGPT-4 to identify specific phrases or sections in the articles that justify the rating to each of the 14 criteria in the prompt. We provide four examples below.

Example 1: Mudrick Capital Management article in the WSJ (June 11, 2021)

1. (Score: 6) The article provides specific information about Mudrick's trading strategies, derivative positions, and market outcomes, e.g., "sold call options, derivative contracts meant to hedge the fund's exposure to AMC should the stock price founder."
2. (Score: 2) The article is mostly detailed, leaving little room for vagueness, but does not provide specific details on the exact financial figures of losses except in aggregate percentages.
3. (Score: 4) References to specific events like "Reddit-fueled buying" and descriptions of high-stakes trading make the content engaging.
4. (Score: 6) The structured presentation of Mudrick's strategies and outcomes, such as "Mudrick Capital's losses highlight how risky retaining significant exposure to such companies can be."
5. (Score: 3) Neutral to mildly critical tone, with only brief mentions of profitable maneuvers, e.g., "Mudrick immediately sold the stock at a profit."
6. (Score: 3) Mentions of losses and setbacks, like "the winning trade had turned into a bust."
7. (Score: 4) The depth of analysis about hedge fund strategies shows an interest, although it's very technical.
8. (Score: 3) Limited emotional language but includes impactful terms like "boomeranged" and "swarming."
9. (Score: 2) The narrative remains factual without supporting the fund's decisions, focusing on description rather than endorsement.
10. (Score: 1) Lacks a direct opposing stance, mainly provides a narrative of events.
11. (Score: 1) Sticks to reporting the events and strategies factually.
12. (Score: 1) The article does not align or agree with external opinions but reports on varied perspectives.
13. (Score: 1) Does not explicitly disagree, maintaining an objective reporting tone.
14. (Score: 7) Strictly factual and analytical, without expressing views on others' opinions.

Example 2: Retelling of Mudrick Capital Management retelling article in the NYP (June 14, 2021)

1. (Score: 5) Provides an overview of the situation but with less detail than the first, e.g., "a series of bets it had also made against the company."
2. (Score: 3) More summarizing and less detail on specific strategies and outcomes than the first article.
3. (Score: 4) Maintains interest with dramatic descriptions of trading outcomes, such as "costing the fund hundreds of millions of dollars."
4. (Score: 5) Clear and concise, though not as richly detailed as the first, making it slightly less 'well-written' in comparison.
5. (Score: 3) Neutral to slightly critical, focusing on the fund's decisions without endorsing them.
6. (Score: 3) Highlights losses which portray a negative situation, e.g., "tanked its value by 10 percent."
7. (Score: 4) Engages the reader with the narrative of loss and market dynamics, albeit less detailed.
8. (Score: 3) Similar to the first, maintains a professional tone with limited emotional content.
9. (Score: 1) There's no evident support, focusing on describing events.
10. (Score: 2) Slightly more critical by underscoring the mismanagement and outcomes of the bets.
11. (Score: 2) Primarily factual but the slightly critical tone suggests a minor departure from neutrality.
12. (Score: 1) Does not show agreement with external opinions, focusing on reporting.
13. (Score: 1) No explicit disagreements presented.
14. (Score: 6) Remains mostly neutral, slightly less so compared to the first article.

Example 3: PepsiCo article in the WSJ (December 5, 2022

1. (Score: 6) "Hundreds of jobs will be eliminated... based in Purchase, N.Y., and its North America snacks and packaged-foods business, which has headquarters in Chicago and Plano, Texas".
2. (Score: 2) Specificity is high, with minimal vagueness, providing clear information on locations and divisions affected.
3. (Score: 4) The report of layoffs extends beyond tech and media, potentially engaging a broader audience.
4. (Score: 6) Clear, professional language is used, such as "simplify the organization so we can operate more efficiently".
5. (Score: 3) Neutral, focusing on the fact of layoffs without positive spin.
6. (Score: 3) Discusses job cuts, which is inherently a bit negative, but stays mostly neutral.
7. (Score: 4) The detail provided shows a focused interest in the business decisions of PepsiCo.
8. (Score: 3) Maintains a professional tone with minimal emotional language.
9. (Score: 2) Largely factual reporting without showing support for the layoffs.
10. (Score: 1) No opposition expressed, purely factual.
11. (Score: 1) The article maintains journalistic neutrality.
12. (Score: 1) No reference to other opinions.
13. (Score: 1) No disagreement presented.
14. (Score: 7) Focuses strictly on reporting the facts without engaging with external opinions.

Example 4: PepsiCo retelling article in the WP (December 6, 2022).

1. (Score: 5) While detailed, it focuses more on broader economic implications, “If a company performing well... is reducing its workforce then it is likely others will be considering similar moves”.
2. (Score: 3) Broader discussions about the economy, like “headed toward a potentially pretty significant economic slowdown”.
3. (Score: 4) Engages readers with analysis on economic trends and implications of corporate decisions.
4. (Score: 5) Good analysis but slightly less direct, “This, rather than inflation, could become the major economic problem in 2023”.
5. (Score: 3) Neutral to slightly negative, similar to the first article.
6. (Score: 3) Reporting on layoffs with a slightly more analytical perspective on economic downturns.
7. (Score: 4) Detailed economic analysis shows deep interest in broader economic trends.
8. (Score: 3) Remains analytical and professional, with a focus on economic implications rather than emotional response.
9. (Score: 1) No support indicated, maintains a critical analysis.
10. (Score: 2) Offers a critical view of economic trends and potential downturns.
11. (Score: 2) Slightly more opinionated in its economic analysis than the first article.
12. (Score: 1) References expert opinions but does not explicitly agree.
13. (Score: 1) Provides analysis but does not explicitly disagree.
14. (Score: 6) While incorporating expert insights, it remains mostly neutral and analytical.

Appendix E. Definitions of variables

Variable	Definition
$retold$	An indicator variable for whether the exclusive WSJ article has been retold by other news media.
$retelling$	An indicator variable for whether the article is a retelling of an exclusive WSJ article.
$NegPer$	The difference between the negative personalization scores between the original story and the retelling. The negative personalization score for each article (the exclusive WSJ article and the corresponding retelling articles) is calculated as: $Neg - Pos + Opinion - Fact - Appeal$.
$NewsOrgFreq$	The number of retelling articles of exclusive WSJ articles by a news media organization in the past. The non-retelling articles have 0 for this variable.
$DaysBetween$	The days between the exclusive WSJ article and the retelling article. The non-retelling articles have 0 for this variable.
$RetellingsSameDay$	The number of retelling articles of the same exclusive WSJ article on the same day. The non-retelling articles have 0 for this variable.
$Opinion - Fact$	The difference between *Opinion* and $Fact$. We sum ratings on elements 8, 11 (reverse-coded), and 14 (reverse-coded) in the prompt as our *Opinion* variable. The sum of the ratings on the first two elements in the prompt is our $Fact$ variable. This variable captures the level of opinions less level of factual details in the text. The non-retelling articles have 0 for this variable.
$Neg - Pos$	The difference between Neg and Pos. Neg is the sum of elements 6, 10, and 13 in the prompt. Pos is the sum of elements 5, 9, and 12 in the prompt. Negativity or Neg (positivity or Pos) is the degree of opposition (support) or disagreement (agreement) conveyed in the text. This variable captures the net pessimism in the text. The non-retelling articles have 0 for this variable.
Appeal	The sum of elements 4, 5, and 7 in the prompt. This variable captures how appealing the text is perceived. The non-retelling articles have 0 for this variable.
$abret_{t+1,t+21}$	The compounded abnormal returns between t and $t + 21$ adjusted using the Fama-French three-factor model plus momentum (Fama and French, 1993; Carhart, 1997). We estimate expected returns using a 100-day trading window with a 50-day gap (70 day minimum window).
$abret_{t-19,t}$	The compounded abnormal returns between $t - 19$ and t adjusted using the Fama-French three-factor model plus momentum (Fama and French, 1993; Carhart, 1997). We estimate expected returns using a 100-day trading window with a 50-day gap (70 day minimum window).

Variable	**Definition**
$AbLogTurnover_{t+1,t+21}$	The difference between average daily log turnover between $t-62$ and t and the average daily log turnover between $t-140$ and $t-20$ (6-month period, skipping the most recent month). Turnover is calculated as the daily volume scaled by shares outstanding.
$AbLogTurnoverInst_{t+1,t+63}$	The difference between average daily log turnover for institutional traders between $t+1$ and $t+63$ and the average daily log turnover for institutional traders between $t-252$ and $t-63$ (9-month period, skipping the most recent quarter). Turnover is calculated as the daily volume scaled by shares outstanding.
$AbLogTurnoverRetail_{t+1,t+63}$	The difference between average daily log turnover for retail traders between $t+1$ and $t+63$ and the average daily log turnover for retail traders between $t-252$ and $t-63$ (9-month period, skipping the most recent quarter). Turnover is calculated as the daily volume scaled by shares outstanding.
$AnalystDisper_{t+1,t+63}$	The standard deviation of analyst forecasts made in the quarter post article publication (between $t+1$ and $t+63$).
$ImpliedVolCall_{t+1,t+21}$	The average implied volatility between $t+1$ and $t+21$ for near the money call options (0.95 and 1.05 moneyness). We include call options with a minimum of 7 days till expiration and a maximum of 186 days till expiration. We weight the average implied volatility across contracts by open interest.
$ImpliedVolPut_{t+1,t+21}$	The average implied volatility between $t+1$ and $t+21$ for near the money put options (0.95 and 1.05 moneyness). We include put options with a minimum of 7 days till expiration and a maximum of 186 days till expiration. We weight the average implied volatility across contracts by open interest.
$Dis_{t+1,t+21}$	The disagreement among investors between $t+1$ and $t+21$. Disagreement is defined as the standard deviation of the binary sentiment measure on social media posts from StockTwits (-1 for bearish, 1 for bullish). The data is from Cookson and Niessner (2020).
$TickerCount$	The number of tickers mentioned in the exclusive WSJ article.
$readability$	The Flesch Reading Ease of the text to capture linguistic complexity (Kincaid et al., 1975).
$neg-pos_{lm}$	Is the net of the proportion of negative words in the text and positive words in the text according to the 2018 version of the Loughran and McDonald (2011) lexicon.
$numbers$	The total instances of digits or numbers identified in the text.
$complex$	The percentage of words in the text that are complex. We define complex words as those with two more syllables.

Variable	**Definition**
beta	Market beta is estimated using the Fama and MacBeth (1973) regressions using weekly returns and equal weighted market returns for three years.
idiovol	Idiosyncratic volatility is calculated following Ali et al. (2003) by taking the standard deviation of residuals of weekly returns on weekly equal-weighted market returns for the three-year window.
ill	Amihud (2002) illiquidity measure is calculated by taking the average of daily absolute returns divided by dollar volume.
lev	Leverage ratio is calculated following Bhandari (1988) by taking total liabilities and dividing by the market capitalization.
roic	The profitability of companies is calculated using the definition in Brown and Rowe (2007) by taking the annual earnings before interest and taxes net of non-operating income and scale by non-cash enterprise value.
bm	The book-to-market ratio is calculated following Rosenberg et al. (1985) as the book value of equity divided by market capitalization.
$log(mv)$	The natural logarithm of market capitalization of the stock calculated by multiplying the price of the share by the shares outstanding.

Appendix F. Determinants of retelling articles

This table presents coefficients from a stock-article level ordinary least squares regressions of *retelling* dummy on article and stock characteristics. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Dept Variable	(1) *retelling*
$log(TickerCount)$	0.0251***
	(3.29)
$readability$	-0.0102***
	(-15.52)
$neg - pos_{lm}$	0.0022
	(1.31)
$numbers$	-0.0031***
	(-8.04)
$complex$	-2.1424***
	(-13.51)
$beta$	-0.0143**
	(-2.17)
$idiovol$	0.2432
	(1.49)
ill	-0.0085
	(-0.45)
lev	-0.0005
	(-0.66)
$roic$	0.0031
	(0.53)
bm	-0.0235***
	(-3.51)
$log(mv)$	0.0015
	(0.90)
$abret_{t-19,t}$	0.0210**
	(2.44)
_cons	1.0298***
	(14.03)
N	16768
adj. R-sq	0.125

References

Ali, A., Hwang, L.-S., Trombley, M. A., 2003. Arbitrage risk and the book-to-market anomaly. Journal of Financial Economics 69, 355–373.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., 2018. The economic effects of social networks: Evidence from the housing market. Journal of Political Economy 126, 2224–2276.

Barasch, A., Berger, J., 2014. Broadcasting and narrowcasting: How audience size affects what people share. Journal of Marketing Research 51, 286–299.

Barber, B. M., Huang, X., Odean, T., Schwarz, C., 2022. Attention-induced trading and returns: Evidence from robinhood users. Journal of Finance 77, 3141–3190.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., Odean, T., 2009. Just how much do individual investors lose by trading? Review of Financial Studies 22, 609–632.

Barber, B. M., Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. Journal of Finance 55, 773–806.

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

Barber, B. M., Odean, T., Zhu, N., 2008. Do retail trades move markets? Review of Financial Studies 22, 151–186.

Baron, D. P., 2005. Competing for the public through the news media. Journal of Economics & Management Strategy 14, 339–376.

Barrett, J., Nyhof, M., 2001. Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. Journal of Cognition and Culture 1, 69–100.

Besley, T., Prat, A., 2006. Handcuffs for the grabbing hand? media capture and government accountability. American Economic Review 96, 720–736.

Bhandari, L. C., 1988. Debt/equity ratio and expected common stock returns: Empirical evidence. Journal of Finance 43, 507–528.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. Journal of Finance 76, 2249–2305.

Brown, D. P., Rowe, B., 2007. The productivity premium in equity returns. Available at SSRN 993467 .

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

Chen, Y., Kelly, B. T., Xiu, D., 2023. Expected returns and large language models. Available at SSRN 4416687 .

Cookson, J. A., Engelberg, J. E., Mullins, W., 2023. Echo chambers. Review of Financial Studies 36, 450–500.

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

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 .

Diether, K. B., Malloy, C. J., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. Journal of Finance 57, 2113–2141.

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

Engelberg, J., Sasseville, C., Williams, J., 2012. Market madness? the case of mad money. Management Science 58, 351–364.

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

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

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. Journal of Finance 64, 2023–2052.

Garcia, D., 2013. Sentiment during recessions. Journal of Finance 68, 1267–1300.

Gentzkow, M., Shapiro, J. M., 2006. Media bias and reputation. Journal of Political Economy 114, 280–316.

Gentzkow, M., Shapiro, J. M., 2008. Competition and truth in the market for news. Journal of Economic perspectives 22, 133–154.

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

Goldman, E., Gupta, N., Israelsen, R., 2024. Political polarization in financial news. Journal of Financial Economics 155, 103816.

Gurun, U. G., Butler, A. W., 2012. Don't believe the hype: Local media slant, local advertising, and firm value. Journal of Finance 67, 561–598.

Hansen, A. L., Kazinnik, S., 2023. Can chatgpt decipher fedspeak? Working Paper .

Hillert, A., Jacobs, H., Müller, S., 2014. Media makes momentum. Review of Financial Studies 27, 3467–3501.

Hribar, P., McInnis, J., 2012. Investor sentiment and analysts' earnings forecast errors. Management science 58, 293–307.

Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. Journal of Finance 56, 387–396.

Hyman, I. E., 1994. Conversational remembering: Story recall with a peer versus for an experimenter. Applied Cognitive Psychology 8, 49–66.

Jaccard, P., 1901. Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines. Bull Soc Vaudoise Sci Nat 37, 241–272.

Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. Journal of Financial Economics 132, 126–149.

Jones, C. M., Lamont, O. A., 2002. Short-sale constraints and stock returns. Journal of Financial Economics 66, 207–239, limits on Arbitrage.

Karpoff, J. M., 1986. A theory of trading volume. The Journal of Finance 41, 1069–1087.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., Chissom, B. S., 1975. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel .

Lopez-Lira, A., Tang, Y., 2023. Can chatgpt forecast stock price movements? return predictability and large language models. Working Paper .

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. Journal of Finance 66, 35–65.

Melumad, S., Meyer, R., Kim, Y. D., 2021. The dynamics of distortion: How successive summarization alters the retelling of news. Journal of Marketing Research 58, 1058–1078.

Mullainathan, S., Shleifer, A., 2005. The market for news. American economic review 95, 1031–1053.

Obaid, K., Pukthuanthong, K., 2022. A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. Journal of Financial Economics 144, 273–297.

Reuter, J., Zitzewitz, E., 2006. Do ads influence editors? advertising and bias in the financial media. The Quarterly Journal of Economics 121, 197–227.

Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. Journal of Portfolio Management 11, 9–16.

Solomon, D. H., Soltes, E., Sosyura, D., 2014. Winners in the spotlight: Media coverage of fund holdings as a driver of flows. Journal of Financial Economics 113, 53–72.

Stapert, J. C., Clore, G. L., 1969. Attraction and disagreement-produced arousal. Journal of Personality and Social Psychology 13, 64.

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

Tetlock, P. C., 2010. Does public financial news resolve asymmetric information? Review of Financial Studies 23, 3520–3557.

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

Wiggers, K., 2023. Openai releases gpt-4, a multimodal ai that it claims is state-of-the-art. Accessed on April 2, 2024.

Yang, S., 2023. Predictive patentomics: Forecasting innovation success and valuation with chatgpt. Working Paper .

Table 1: Exclusive WSJ articles sample

This table displays descriptive information for the exclusive WSJ articles sample. The sample consists of all the WSJ articles labelled as exclusive (‘PMDM’) directly by Dow Jones. The sample period is between 2013 and 2022.

	N				
# of retold articles-match observations	1,940				
# of unique retold articles	1,351				
# of non-retold articles	27,524				
# of retold articles by year	2013: 100	2014: 174	2015: 174	2016: 121	2017: 143
	2018: 150	2019: 132	2020: 137	2021: 113	2022: 107
# of non-retold articles by year	2013: 4957	2014: 6485	2015: 4396	2016: 1852	2017: 1448
	2018: 1484	2019: 1772	2020: 1741	2021: 1753	2022: 1636
# of retold articles with tickers	359				
# of non-retold articles with tickers	9,388				
# of news organizations	18				
News organizations with most retellings	INVD’: 589	’NYPO’: 356	’WPCO’: 226	’USAT’: 212	’NYDN’: 98
	mean	median	std	max	min
# of tickers reflected in retold articles	2.702	2	3.531	42	1
# of tickers reflected in non-retold articles	1.661	1	1.38	30	1
# of retellings per article	1.404	1	0.818	5	1
# of days between and retellings	2.635	1	3.38	14	0

Table 2: ChatGPT-4 distortion variables

Panel A reports the descriptive statistics for the distortion measures derived from the exclusive WSJ articles that have been retold (A) and the respective retelling articles ($Match$) utilizing ChatGPT. Panel B presents the mean difference for each of the distortion measures across the exclusive WSJ article sample and the respective retelling articles sample. The sample is between 2013 and 2022. All variables are defined in Appendix E.

Panel A: Distortion Variables (ChatGPT)						
Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
$Fact_A$	1,940	11.915	12.000	1.141	14.000	3.000
$Fact_{Match}$	1,940	10.458	11.000	2.697	14.000	2.000
$Opinion_A$	1,940	8.728	8.000	3.129	20.000	3.000
$Opinion_{Match}$	1,940	9.159	9.000	3.204	21.000	3.000
Neg_A	1,940	6.439	6.000	2.465	21.000	3.000
Neg_{Match}	1,940	7.103	7.000	3.097	21.000	3.000
Pos_A	1,940	6.096	6.000	2.064	15.000	3.000
Pos_{Match}	1,940	5.978	6.000	2.107	17.000	3.000
$Appeal_A$	1,940	15.677	16.000	2.026	21.000	4.000
$Appeal_{Match}$	1,940	14.809	15.000	2.990	21.000	4.000
Panel B: Means Difference						
Variable	Diff (A-Match)	t-statistic	p-value			
Fact	1.457	21.910	0.000			
$Opinion$	-0.431	-4.243	0.000			
Neg	-0.664	-7.388	0.000			
Pos	0.119	1.771	0.077			
$Appeal$	0.868	10.585	0.000			

Table 3: Textual variables

This table displays the descriptive statistics for various textual variables derived from the exclusive WSJ articles that have been retold (A) and the respective retelling articles ($Match$). *readability* is the Flesch Reading Ease designed to assess the clarity of English writing. *numbers* is the count of numbers in the text. *complex* is the percentage of words in the text that have two or more syllables. $neg - pos_{lm}$ is the percentage of negative words minus positive words using the 2018 version of the Loughran and McDonald (2011) lexicon. The sample is between 2013 and 2022. All variables are defined in Appendix E.

Automated Text Variables						
Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
$readability_A$	1,940	57.225	57.060	8.301	84.170	15.650
$readability_{Match}$	1,940	51.108	51.565	12.872	98.410	-9.820
$numbers_A$	1,940	14.814	13.000	8.256	69.000	6.000
$numbers_{Match}$	1,940	9.880	6.000	11.548	187.000	0.000
$complex_A$	1,940	0.170	0.169	0.035	0.312	0.050
$complex_{Match}$	1,940	0.162	0.162	0.037	0.321	0.039
$neg - pos_A$	1,940	1.844	1.373	2.088	19.231	-2.308
$neg - pos_{Match}$	1,940	1.863	1.404	2.286	14.894	-4.930
$cosine\_similarity$	1,940	0.704	0.721	0.114	0.929	0.178

Table 4: Firm characteristics and market variables

Panel A reports descriptive statistics for the characteristics of firms in the exclusive WSJ articles. Panel B presents the descriptive statistics for the market variables of firms in the exclusive WSJ articles. Panels C and D focus on comparing the group of variables in Panels A and B, respectively, for exclusive WSJ articles that are *retold* (*retold*1) and those that are not retold (*retold*0). The sample is between 2013 and 2022. All variables are defined in Appendix E.

Panel A: Firm Characteristics

Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
retelling	16,863	0.071	0.000	0.257	1.000	0.000
retold	16,863	0.061	0.000	0.240	1.000	0.000
beta	16,819	1.118	1.072	0.454	3.325	-0.124
idiovol	16,819	0.035	0.029	0.019	0.260	0.010
ill	16,863	0.002	0.000	0.037	2.205	0.000
lev	16,863	2.807	0.706	4.723	59.307	0.000
roic	16,819	0.095	0.081	0.296	1.090	-12.468
bm	16,863	0.481	0.333	0.474	4.008	-1.212
mve_m	16,863	139028.347	60530.352	248767.872	2902368.097	7.877

Panel B: Returns

Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
$ret_{t-1,t}$	16,863	0.006	0.002	0.096	7.816	-0.609
$ret_{t+1,t+21}$	16,862	0.012	0.012	0.111	3.392	-0.850
$abret_{t+1,t+21}$	16,862	-0.002	-0.003	0.105	2.860	-0.942
$AbLogTurnover_{t+1,t+21}$	16,855	0.001	0.000	0.015	0.587	-0.214

Panel C: Firm Characteristics by Retelling						
Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
beta_retold0	15,791	1.125	1.080	0.455	3.325	-0.124
beta_retold1	1,028	1.000	0.950	0.414	3.140	-0.109
idiovol_retold0	15,791	0.035	0.029	0.019	0.260	0.010
idiovol_retold1	1,028	0.034	0.029	0.020	0.260	0.014
ill_retold0	15,830	0.002	0.000	0.038	2.205	0.000
ill_retold1	1,033	0.001	0.000	0.015	0.348	0.000
lev_retold0	15,830	2.890	0.723	4.812	59.307	0.000
lev_retold1	1,033	1.535	0.558	2.769	22.025	0.000
roic_retold0	15,787	0.095	0.079	0.300	1.090	-12.468
roic_retold1	1,032	0.104	0.098	0.227	0.767	-2.547
bm_retold0	15,830	0.490	0.340	0.481	4.008	-1.212
bm_retold1	1,033	0.342	0.276	0.325	2.477	-0.879
mve_m_retold0	15,830	136427.678	59605.010	246043.314	2902368.097	7.877
mve_m_retold1	1,033	178881.769	84148.665	284481.774	2509775.134	27.040

Panel D: Returns by Retelling						
Variable	Number of Obs	Mean	Median	Std Dev	Max	Min
$ret\_retold0_{t+1,t+21}$	15,829	0.012	0.012	0.110	3.392	-0.780
$ret\_retold1_{t+1,t+21}$	1,033	0.004	0.007	0.126	1.170	-0.850
$abret\_retold0_{t+1,t+21}$	15,829	-0.002	-0.002	0.104	2.860	-0.901
$abret\_retold1_{t+1,t+21}$	1,033	-0.012	-0.011	0.123	1.096	-0.942
$AbLogTurnover\_retold0_{t+1,t+21}$	15,825	0.001	0.000	0.015	0.587	-0.214
$AbLogTurnover\_retold1_{t+1,t+21}$	1,030	0.004	0.000	0.022	0.335	-0.168

Panel E: Means Difference			
Variable	Diff (retold1-retold0)	t-statistic	p-value
beta	-0.125	-9.334	0.000
idiovol	0.000	0.048	0.962
ill	-0.001	-1.249	0.212
lev	-1.354	-14.366	0.000
roic	0.009	1.238	0.216
bm	-0.148	-13.701	0.000
mve_m	42454.091	4.683	0.000
$ret_{t+1,t+21}$	-0.008	-1.994	0.046
$abret_{t+1,t+21}$	-0.011	-2.747	0.006
$AbLogTurnover_{t+1,t+21}$	0.003	3.745	0.000

Table 5: Retelling articles and distortion

This table presents coefficients from article level ordinary least squares regressions of ChatGPT-4 distortion measures on article characteristics. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include month-year fixed effects and the t-statistics are based on standard errors clustered by article. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Dept Variable	(1) $Opinion - Fact$	(2) $Opinion - Fact$	(3) $Neg - Pos$	(4) $Neg - Pos$	(5) $log(Appeal)$	(6) $log(Appeal)$
$retelling$	1.9146***	1.9514***	0.7986***	0.8488***	-0.0656***	-0.0533***
	(28.82)	(20.77)	(10.48)	(8.73)	(-10.97)	(-6.33)
$log(TickerCount)$		-0.7566***		-0.6240***		-0.0152***
		(-5.79)		(-5.22)		(-2.61)
$readability$		0.0252***		0.0013		-0.0010**
		(3.76)		(0.22)		(-2.25)
$neg - pos_{lm}$		0.1301***		0.5099***		-0.0093***
		(4.20)		(15.92)		(-5.18)
$numbers$		-0.0284***		0.0030		0.0036***
		(-4.52)		(0.56)		(5.92)
$complex$		3.2238		4.2538**		0.0267
		(1.46)		(2.16)		(0.20)
_cons	-3.2182***	-4.4181***	0.2820***	-0.9951*	2.8045***	2.8323***
	(-44.37)	(-6.24)	(3.61)	(-1.66)	(817.17)	(60.43)
N	3794	3794	3794	3794	3794	3794
adj. R-sq	0.111	0.150	0.094	0.211	0.049	0.107

Table 6: The role of specialization, memory, and motivation on distortion

This table presents coefficients from article level ordinary least squares regressions of ChatGPT-4 distortion measures on article characteristics. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include month-year fixed effects and the t-statistics are based on standard errors clustered by article. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Dept Variable	(1) $Opinion - Fact$	(2) $Neg - Pos$	(3) $log(Appeal)$
$log(NewsOrgFreq)$	0.2926***	-0.0750*	-0.0189***
	(6.81)	(-1.70)	(-6.65)
$log(DaysBetween)$	0.3454***	0.5110***	0.0070
	(3.54)	(4.44)	(1.25)
$log(RetellingsSameDay)$	0.1570	0.6224***	0.0241**
	(0.84)	(3.37)	(2.19)
$log(TickerCount)$	-0.7631***	-0.6144***	-0.0146**
	(-5.87)	(-5.18)	(-2.56)
$readability$	0.0257***	-0.0046	-0.0014***
	(3.65)	(-0.72)	(-2.93)
$neg - pos_{lm}$	0.1310***	0.5066***	-0.0094***
	(4.19)	(15.77)	(-5.31)
$numbers$	-0.0329***	0.0030	0.0038***
	(-5.02)	(0.57)	(6.22)
$complex$	3.1124	3.6203*	-0.0236
	(1.38)	(1.80)	(-0.18)
_cons	-4.2814***	-0.5389	2.8610***
	(-5.81)	(-0.85)	(57.45)
N	3794	3794	3794
adj. R-sq	0.143	0.218	0.114

Table 7: Abnormal returns and distortion

This table presents coefficients from a stock-article level ordinary least squares regressions of contemporaneous and future abnormal performance on article characteristics, including the ChatGPT-4 distortion variables and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dept Variable	$log(abret_{t-19,t})$		$log(abret_{t+1,t+21})$			
$retelling$	0.0090	0.0090	-0.0099**	-0.0102***	-0.0107***	
	(1.45)	(1.41)	(-2.48)	(-2.60)	(-2.58)	
$NegPer$						-0.0013**
						(-2.42)
$log(NewsOrgFreq)$						-0.0017
						(-1.09)
$log(DaysBetween)$						-0.0077
						(-1.60)
$log(RetellingsSameDay)$						-0.0005
						(-0.06)
$log(TickerCount)$		-0.0030			-0.0010	-0.0011
		(-1.14)			(-0.48)	(-0.54)
$readability$		0.0001			-0.0001	-0.0001
		(0.55)			(-0.84)	(-1.11)
$neg-pos_{lm}$		-0.0043***			-0.0004	-0.0004
		(-5.86)			(-0.60)	(-0.66)
$numbers$		-0.0002**			-0.0000	0.0000
		(-2.40)			(-0.10)	(0.15)
$complex$		0.0116			0.0128	0.0070
		(0.31)			(0.40)	(0.21)
$beta$	0.0032	0.0037		-0.0051	-0.0050	-0.0052
	(0.28)	(0.32)		(-0.68)	(-0.66)	(-0.69)
$idiovol$	-0.7068*	-0.7034*		-0.9492***	-0.9536***	-0.9540***
	(-1.84)	(-1.84)		(-4.74)	(-4.77)	(-4.78)
ill	0.0853	0.0862		-0.0306	-0.0310	-0.0310
	(0.91)	(0.93)		(-0.42)	(-0.42)	(-0.42)
lev	-0.0004	-0.0003		-0.0005	-0.0005	-0.0005
	(-0.61)	(-0.50)		(-0.90)	(-0.87)	(-0.86)
$roic$	-0.0383*	-0.0375*		-0.0078	-0.0077	-0.0078
	(-1.89)	(-1.85)		(-0.90)	(-0.89)	(-0.90)
bm	0.0084	0.0087		0.0042	0.0043	0.0042
	(0.68)	(0.71)		(0.68)	(0.71)	(0.68)
$log(mv)$	-0.0063***	-0.0063***		-0.0037***	-0.0038***	-0.0039***
	(-4.20)	(-4.19)		(-3.43)	(-3.48)	(-3.53)
$retold$	0.0118*	0.0120*	-0.0141***	-0.0144***	-0.0144***	-0.0144***
	(1.67)	(1.70)	(-2.59)	(-2.65)	(-2.67)	(-2.67)
$abret_{t-19,t}$			-0.0290	-0.0268	-0.0269	-0.0269
			(-0.97)	(-0.89)	(-0.89)	(-0.89)
_cons	0.1310***	0.1349***	-0.0059***	0.0987***	0.1055***	0.1101***
	(4.32)	(3.89)	(-5.16)	(4.45)	(4.13)	(4.26)
N	16768	16768	16855	16767	16767	16767
adj. R-sq	0.028	0.032	0.073	0.094	0.094	0.094

Table 8: Abnormal turnover and distortion

This table presents coefficients from a stock-article level ordinary least squares regressions of future abnormal turnover on article characteristics including the ChatGPT-4 distortion variables and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Dept Variable	(1)	(2) $AbLogTurnover_{t+1,t+21}$	(3)	(4)
$retelling$	0.0012***	0.0015***	0.0014***	
	(2.60)	(3.16)	(2.80)	
$NegPer$				0.0002***
				(4.16)
$log(NewsOrgFreq)$				0.0006***
				(2.75)
$log(DaysBetween)$				-0.0001
				(-0.12)
$log(RetellingsSameDay)$				-0.0000
				(-0.04)
$log(TickerCount)$			-0.0006***	-0.0006**
			(-2.59)	(-2.28)
$readability$			0.0000	0.0000
			(0.17)	(1.08)
$neg-pos\_lm$			0.0002***	0.0002***
			(3.27)	(3.44)
$numbers$			-0.0000	-0.0000
			(-0.76)	(-1.38)
$complex$			0.0007	0.0028
			(0.16)	(0.63)
$beta$		0.0005	0.0005	0.0005
		(0.64)	(0.65)	(0.69)
$idiovol$		-0.0016	-0.0034	-0.0036
		(-0.03)	(-0.07)	(-0.08)
ill		0.0059	0.0058	0.0057
		(0.83)	(0.82)	(0.82)
lev		0.0003**	0.0003**	0.0003**
		(2.16)	(2.18)	(2.17)
$roic$		-0.0001	-0.0001	-0.0001
		(-0.05)	(-0.06)	(-0.05)
bm		-0.0016	-0.0016	-0.0016
		(-1.31)	(-1.34)	(-1.34)
$log(mv)$		-0.0009***	-0.0009***	-0.0009***
		(-5.11)	(-5.06)	(-5.03)
$retold$	0.0021***	0.0024***	0.0023***	0.0023***
	(3.63)	(4.13)	(3.95)	(3.95)
$abret_{t-19,t}$	0.0214***	0.0208***	0.0208***	0.0208***
	(2.99)	(2.95)	(2.96)	(2.95)
_cons	0.0008***	0.0163***	0.0168***	0.0153***
	(5.94)	(4.18)	(3.29)	(3.01)
N	16848	16760	16760	16760
adj. R-sq	0.159	0.171	0.172	0.173

Table 9: Abnormal turnover by retail and institutional investors

This table presents coefficients from a stock-article level ordinary least squares regressions of future abnormal turnover for retail and institutional traders on article characteristics and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)
	$AbLogTurnover_{t+1,t+63}$	
Dept Variable	Retail	Institutional
$retelling$	0.0505**	0.1125
	(2.08)	(0.54)
$log(TickerCount)$	-0.0191**	0.0519
	(-2.29)	(0.41)
$readability$	-0.0005	0.0135*
	(-0.72)	(1.68)
$neg - pos_{lm}$	0.0131***	-0.0832***
	(3.95)	(-2.70)
$numbers$	0.0001	0.0011
	(0.30)	(0.27)
$complex$	0.0461	4.9912**
	(0.29)	(2.08)
$beta$	0.0255	-0.7940**
	(0.68)	(-2.35)
$idiovol$	0.5068	1.9009
	(0.31)	(0.24)
ill	0.3344	1.2184
	(1.19)	(1.59)
lev	0.0085**	0.0339
	(2.00)	(1.13)
$roic$	0.0297	0.0146
	(0.64)	(0.05)
bm	-0.0294	-0.0776
	(-0.73)	(-0.27)
$log_m v$	-0.0194***	0.0608
	(-2.77)	(0.94)
$retold$	0.0667***	-0.1994
	(2.59)	(-1.01)
$abret_{t-19,t}$	0.7104***	0.0120
	(3.37)	(0.01)
_cons	0.3194*	-1.9684
	(1.70)	(-1.22)
N	15357	14431
adj. R-sq	0.252	0.012

Table 10: Analyst forecast dispersion

This table presents coefficients from a stock-article level ordinary least squares regressions of future analyst forecast dispersion on article characteristics and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)
Dept Variable	$log(AnalystDisper_{t+1,t+63})$		
$retelling$	0.2442***	0.2775***	0.2930***
	(4.36)	(5.12)	(5.14)
$log(TickerCount)$			-0.0358
			(-1.15)
$readability$			0.0000
			(0.03)
$neg-pos_{lm}$			0.0029
			(0.24)
$numbers$			0.0012
			(1.14)
$complex$			0.7147
			(1.41)
$beta$		0.2916***	0.2915***
		(4.69)	(4.75)
$idiovol$		6.5722***	6.5084***
		(3.10)	(3.08)
ill		-0.4791	-0.4882
		(-1.16)	(-1.17)
lev		-0.0292***	-0.0299***
		(-2.81)	(-2.88)
$roic$		-0.0061	-0.0048
		(-0.09)	(-0.07)
bm		-0.1129	-0.1110
		(-1.40)	(-1.38)
$log(mv)$		-0.0949***	-0.0962***
		(-5.41)	(-5.47)
$retold$	0.2535***	0.2745***	0.2777***
	(4.48)	(5.15)	(5.17)
$abret_{t-19,t}$	0.2385***	0.1925***	0.1916***
	(5.80)	(2.68)	(2.68)
_cons	-0.3011***	0.8968**	0.8132*
	(-9.49)	(2.37)	(1.75)
N	7287	7255	7255
adj. R-sq	0.071	0.140	0.140

Table 11: Social media disagreement

This table presents coefficients from a stock-article level ordinary least squares regressions of future social media disagreement on article characteristics and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)
		$log(Dis_{t+1,t+21})$	
$retelling$	0.1860***	0.1473***	0.1247***
	(5.96)	(5.57)	(4.83)
$log(TickerCount)$			-0.1407***
			(-6.50)
$readability$			-0.0018**
			(-2.26)
$neg-pos_{lm}$			0.0088**
			(2.42)
$numbers$			-0.0011**
			(-2.19)
$complex$			-0.4252*
			(-1.96)
$beta$		0.1477***	0.1515***
		(2.75)	(2.87)
$idiovol$		17.0449***	16.7924***
		(12.51)	(12.56)
ill		1.0339***	0.9988**
		(2.58)	(2.55)
lev		0.0018	0.0022
		(0.14)	(0.17)
$roic$		-0.0162	-0.0157
		(-0.40)	(-0.39)
bm		0.0526	0.0523
		(0.80)	(0.80)
$log(mv)$		0.3312***	0.3295***
		(22.22)	(22.28)
$retold$	0.1854***	0.1574***	0.1464***
	(5.94)	(5.67)	(5.29)
$abret_{t-19,t}$	-0.0153	0.0624*	0.0572
	(-0.25)	(1.68)	(1.57)
_cons	-0.9665***	-7.6268***	-7.2506***
	(-20.72)	(-25.93)	(-24.51)
N	14495	14421	14421
adj. R-sq	0.230	0.480	0.488

Table 12: Implied volatility

This table presents coefficients from a stock-article level ordinary least squares regressions of future average stock implied volatility on article characteristics and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dept Variable	$log(ImpliedVolCall_{t+1,t+21})$			$log(ImpliedVolPut_{t+1,t+21})$		
$retelling$	0.0450**	0.0556***	0.0499***	0.0450**	0.0556***	0.0499***
	(2.26)	(3.11)	(2.80)	(2.26)	(3.11)	(2.80)
$log(TickerCount)$			-0.0442***			-0.0442***
			(-5.05)			(-5.05)
$readability$			-0.0007			-0.0007
			(-1.59)			(-1.59)
$neg - pos_{lm}$			0.0081***			0.0081***
			(3.38)			(3.38)
$numbers$			-0.0002			-0.0002
			(-0.78)			(-0.78)
$complex$			0.2597**			0.2597**
			(2.15)			(2.15)
$beta$		0.1141***	0.1146***		0.1141***	0.1146***
		(4.04)	(4.11)		(4.04)	(4.11)
$idiovol$		13.1320***	13.0235***		13.1320***	13.0235***
		(14.18)	(14.09)		(14.18)	(14.09)
ill		-3.1232	-3.7401		-3.1232	-3.7401
		(-0.76)	(-0.92)		(-0.76)	(-0.92)
lev		-0.0048	-0.0047		-0.0048	-0.0047
		(-1.58)	(-1.54)		(-1.58)	(-1.54)
$roic$		0.1039**	0.1039**		0.1039**	0.1039**
		(2.23)	(2.24)		(2.23)	(2.24)
bm		-0.0245	-0.0254		-0.0245	-0.0254
		(-0.82)	(-0.85)		(-0.82)	(-0.85)
$log(mv)$		-0.0314***	-0.0334***		-0.0314***	-0.0334***
		(-4.65)	(-5.07)		(-4.65)	(-5.07)
$retold$	0.0732***	0.0773***	0.0725***	0.0732***	0.0773***	0.0725***
	(3.25)	(4.00)	(3.77)	(3.25)	(4.00)	(3.77)
$abret_{t-19,t}$	0.0614	-0.0367	-0.0365	0.0614	-0.0367	-0.0365
	(0.89)	(-0.36)	(-0.37)	(0.89)	(-0.36)	(-0.37)
_cons	-1.3427***	-1.3411***	-1.2600***	-1.3427***	-1.3411***	-1.2600***
	(-78.61)	(-10.18)	(-8.86)	(-78.61)	(-10.18)	(-8.86)
N	15665	15591	15591	15665	15591	15591
adj. R-sq	0.159	0.482	0.487	0.159	0.482	0.487

Table 13: Long-short portfolio analysis

This table reports the average future return or Carhart (1997) alpha (in %) for portfolios sorted on whether stocks are mentioned in a retelling article and the spreads between the retelling and non-retelling portfolios. We sort stocks into two portfolios based on whether they are mentioned in a retelling article at time t and invest in the stocks between $t+1$ and $t+21$. We require portfolios to have at least 10 stocks. The sample is between 2013 and 2022. Newey-West standard errors with five lags are used to compute the t-statistics (reported in brackets). ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Portfolio	*Return*	*CarhartAlpha*	*Market*	*Value*	*Size*	*Momentum*	R2
NoRetelling	0.070***						
Retelling	0.009						
Retelling - *NoRetelling*	-0.061**						
t-statistic	[-2.568]						
NoRetelling		0.011	1.029***	0.146***	0.172***	-0.097***	0.846
Retelling		-0.048	1.061***	0.223***	0.017	-0.163***	0.506
Retelling - *NoRetelling*		-0.059**	0.032	0.077*	-0.155***	-0.066**	0.011
t-statistic		[-2.190]	[1.308]	[1.901]	[-3.479]	[-2.027]	

Table 14: Reversal

This table presents coefficients from a stock-article level ordinary least squares regressions of future abnormal performance on article characteristics, including the ChatGPT-4 distortion variables and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Dept Variable	(1) $log(abret_{t+1,t+21})$	(2) $log(abret_{t+1,t+42})$	(3) $log(abret_{t+1,t+63})$
$retelling$	-0.0107***	-0.0091	-0.0102
	(-2.58)	(-1.57)	(-1.41)
$log(TickerCount)$	-0.0010	-0.0006	0.0003
	(-0.48)	(-0.20)	(0.09)
$readability$	-0.0001	0.0000	-0.0003
	(-0.84)	(0.01)	(-1.28)
$neg - pos_{lm}$	-0.0004	0.0006	0.0013
	(-0.60)	(0.76)	(1.27)
$numbers$	-0.0000	0.0000	0.0001
	(-0.10)	(0.25)	(0.43)
$complex$	0.0128	0.0041	-0.0619
	(0.40)	(0.09)	(-0.97)
$beta$	-0.0050	-0.0075	-0.0011
	(-0.66)	(-0.60)	(-0.06)
$idiovol$	-0.9536***	-1.5697***	-2.4127***
	(-4.77)	(-5.55)	(-5.28)
ill	-0.0310	-0.0130	-0.0747
	(-0.42)	(-0.12)	(-0.68)
lev	-0.0005	-0.0002	-0.0004
	(-0.87)	(-0.21)	(-0.36)
$roic$	-0.0077	-0.0022	-0.0061
	(-0.89)	(-0.19)	(-0.40)
bm	0.0043	0.0019	0.0022
	(0.71)	(0.25)	(0.19)
$log(mv)$	-0.0038***	-0.0038**	-0.0053**
	(-3.48)	(-2.33)	(-2.45)
$retold$	-0.0144***	-0.0123**	-0.0071
	(-2.67)	(-2.00)	(-0.98)
$abret_{t-19,t}$	-0.0269	0.0225	0.0151
	(-0.89)	(0.58)	(0.30)
_cons	0.1055***	0.1159***	0.1875***
	(4.13)	(3.06)	(3.57)
N	16767	16767	16767
adj. R-sq	0.094	0.076	0.068

Table 15: Tone, personalization, and appeal

This table presents coefficients from a stock-article level ordinary least squares regressions of future abnormal performance on article characteristics, including the ChatGPT-4 distortion variables and controls. The sample is between 2013 and 2022. All variables are defined in Appendix E. In all regressions, we include industry fixed effects and the t-statistics are based on standard errors clustered by firm. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

	(1)	(2)	(3)
Dept Variable	$log(abret_{t+1,t+21})$		
$Tone$	-0.0013		
	(-0.96)		
$Personalization$		-0.0026**	
		(-2.40)	
$Appeal$			0.0015*
			(1.85)
$log(NewsOrgFreq)$	-0.0015	-0.0016	-0.0017
	(-0.95)	(-1.02)	(-1.07)
$log(DaysBetween)$	-0.007	-0.0076	-0.0069
	(-1.47)	(-1.61)	(-1.47)
$log(RetellingsSameDay)$	0.0011	-0.0017	0.0013
	(0.14)	(-0.20)	(0.16)
$log(TickerCount)$	-0.0009	-0.0011	-0.0011
	(-0.43)	(-0.53)	(-0.52)
Firm and Article Controls	YES	YES	YES
_cons	0.1066***	0.1095***	0.1083***
	(4.13)	(4.27)	(4.20)
N	16767	16767	16767
adj. R-sq	0.094	0.094	0.094