Viral but Vanishing: Investment Advisors, Social Media, and Regulation

Nazanin Babolmorad Peter Bossaerts Nadia Massoud†

October 2025

ABSTRACT

Over 60% of younger retail investors now rely on peers in investment social media platforms like StockTwits, increasingly substituting professional advice. This paper offers a novel contribution by examining the determinants of investment advice diffusion and the role of professionals in shaping socially sourced narratives. Professionals tend to avoid partisan or emotionally charged content yet assert dominance—being 366% more likely than regular users to post viral content. However, their activity declined sharply after 2018 due to SEC regulatory pressures. These findings suggest that the disintermediation of professional advice arises mainly from regulation rather than platform design or user behavior.

Keywords: Financial Social Networks, Retail Investors, Social Influence, Investment Advisors, Disintermediation of Investment Advice, Government Policy and Regulation, and Regulatory Over-reach.

JEL: D91, G23, G28, G50.

---

† Nazanin Babolmorad is at the Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom; email: nb736@cam.ac.uk. Peter Bossaerts is at the Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom. E-mail address: plb32@cam.ac.uk. Nadia Massoud is at Melbourne Business School, University of Melbourne, Parkvil (VIC), Australia; email: n.massoud@mbs.edu. Bossaerts and Babolmorad gratefully acknowledge funding from the Leverhulme Trust (Leverhulme International Professorship in Neuroeconomics), United Kingdom. Any errors or omissions are the responsibility of the authors. We appreciate the feedback and comments received at seminars held at the University of Cambridge, University of Melbourne, University of Alberta, and Corvinus University. We are especially grateful to Constantine Yannelis, Nandu Nayar, Sylvain Catherine, Paul Irvine, and Tarun Chordia for their valuable advice.

1 Introduction

Traditionally, investment advice has come from professionals, whether directly in person, or more recently, indirectly through robot advice. As such, the financial industry is in control of the narrative, or even the communication channels. However, the landscape has been evolving. According to FINRA (2022), younger retail investors (18–34) trade more frequently, favor individual stocks over mutual funds (79% vs. 58%), prefer online trading (e.g., zero-commission-fee trading platforms) to contacting professionals/firms (62% vs. 44%), and importantly, rely more on social media than professionals for investment insights compared to older investors (60% vs. 35%). This generational shift has introduced the risk of "disintermediation" for the financial industry—particularly for investment advisors— not unlike how commercial banks started to become disintermediated in the 1980s.

Here, we study the percolation of investment advice in financial social networks, and importantly, whether and how professional investment advisors navigate social media and regulation constraints to mitigate the risk of disintermediation posed by social platforms. By focusing on the structure and activity of a prominent investment advice network, StockTwits, we determine (i) how information percolates through the network and what enables users to become influential on StockTwits, (ii) whether and how professionals form a subset of effective influencers, and (iii) to what extent interference by the main professional regulators – the Securities and Exchange Commission (SEC) and the Financial Industry Regulatory Authority (FINRA) – have helped or hindered the industry overcome disintermediation.

Social media platforms such as StockTwits offer advisory services more akin to those of investment advisors. The emphasis is thus on investment advice (i.e., the provision of information on which assets to buy and sell) rather than on comprehensive financial guidance tailored to an individual's broader socio-economic context. The SEC underscores this distinction by identifying *asset allocation advice and equity analysis* as core components of *investment advisory services*, thereby differentiating investment advisors from broader financial planners (see pages 3 and 4 of the Regulation of Investment Advisors). As such, social media increasingly substitutes for traditional advisory services of investment advisors particularly among new generations of investors by offering similar services that may vary in quality. The Congressional Research Service (CRS) refers to this phenomenon as disintermediation (CRS, 2021).

Specifically, our paper studies the role of "peer effects" on social media (DeMarzo, Vayanos and Zwiebel, 2003; Golub and Jackson, 2010), in the percolation and influence of investment advice among new generations of retail investors. We investigate how network design and structure shape these dynamics, and, more importantly, whether and how investment advisors strategically assert their dominance as the primary source of investment guidance. Notably, investment advisors, henceforth referred to as professionals, have not stood still: social media presence has become integral to investment advisors' practices, facilitating retail client engagement, professional networking, and brand expansion (BlackRock, 2024).

One way for professionals to regain control is to strategically secure pivotal positions in online social networks, primarily by accumulating a substantial following and establishing themselves as *potential influencers*. This approach succeeds primarily when the network is designed to value and incentivize professional engagement. It would be difficult, for example, in Reddit/WallStreetBets (WSB), which is marked by user anonymity and support for aggressive trading. There, professionals face a lack of preferential treatment and may even encounter resistance to their advice, as observed during the GameStop short squeeze in early 2021. In contrast, StockTwits is designed to recognize and promote professional input, inspiring its members to actively seek their guidance.

StockTwits' members have the opportunity to garner recognition by formally requesting to be acknowledged as an "Official Account," a designation reserved for professionals or recognized individuals (Fig. A. 1 in the Appendix presents examples of official users). The designation attached to an account is likely to invite many to follow it. However, the effectiveness of social influencers lies in the widespread dissemination of their messages within the network, which may or may not be tied to the number of followers. Professionals thrive in leveraging social media by accumulating a significant follower base, only if the positive correlation between follower count and the intensity of information dissemination within the network identifies them as *effective influencers*.

This disconnect between structure and function may sound surprising: Why would a well-connected network member become well-connected in the first place if it is not because its messages generate interest? *Potential* influencers may not be *effective* influencers in a network for many reasons, such as the network's topology and content curation algorithms. Moreover, a member with a high follower count may become inactive, leading to a decline in influence. A potential influencer might also be followed by other members for reasons that have nothing to do with the content of the messages (in StockTwits: finance). One may decide to follow another node merely out of social curiosity (the node is that of, say, a famous Chief Executive Officer or a politician), socioeconomic affinity (both parent and follower nodes went to the same high school), homophily (both are against financial regulation), or even because StockTwits gives stamps of approval – as mentioned before – by labeling the node as an "official account."

As such, it is compelling to examine the following three questions:

1) What factors determine influence on StockTwits?

2) Do investment advisors effectively leverage these factors to counter the disintermediation threat posed by the network, thereby establishing themselves as influential actors?

3) To what extent do SEC and FINRA regulations constrain professionals' efforts to function as effective social influencers?

This study explores these questions by applying network science, analyzing SEC and FINRA regulations on professionals' use of social media, and utilizing a unique sample of StockTwits network data from January 2014 to December 2020. Specifically, to identify investment advisors on StockTwits, we manually examine the links provided in the Official

Accounts' profiles, such as their business websites and LinkedIn profiles. We also cross-reference their profiles on FINRA BrokerCheck to verify their status as Registered Investment Advisors (hereafter, RIAs).

Additionally, in relation to point (3), the SEC and FINRA govern the conduct of investment advisors, particularly RIAs. Since 2017, both regulatory bodies have progressively intensified their oversight of advisors' social media activity to safeguard retail investors by promoting fair disclosure, preventing regulatory violations, and detecting market manipulation (e.g., pump-and-dump schemes). This scrutiny has further increased since July 2018, resulting in multiple enforcement actions, including regulatory notices, amendments to SEC Form ADV and related rules, and the imposition of fines.[1] As highlighted in K&L Gates' investment management alert, recent enforcement actions raise concerns about unintended consequences. In particular, regulatory overreach may produce a "chilling effect" in which fear of SEC enforcement drives advisors to excessive self-censorship, paralyzing professional insights on social media.

This paper provides a novel empirical analysis to identify the factors that drive the percolation of investment advice and, consequently, influence on financial social networks, with a focus on structural and functional connectivity. **Structural connectivity** refers to follower relationships (who follows whom) identifying *potential* influencers. **Functional connectivity** reflects actual information flow, measured using a cascade model of message resharing. Effective influencers are those whose posts go viral and generate large cascades.

Our analysis reveals that, despite a small proportion of members holding the majority of connections in StockTwits, the expected message dissemination by those well-connected members does not materialize as anticipated. Surprisingly, this implies that a higher follower

[1] This includes but not limited to: Regulatory Notice 16-41 (January 9, 2017), Regulatory Notice 17-18 (April 25, 2017), SEC Form ADV Amendments (October 1, 2017), Administrative proceeding for Violating the Testimonial Rule (July 10, 2018), SEC Amendments to Advertising Rule and Solicitation Rule (November 4, 2019), FINRA Finfluencer Sweep (September, 2021). They all intensify enforcements on FINRA Rules 2210 (Communications with the Public) and 2010 (Standards of Commercial Honor & Principles of Trade).

count does not necessarily correlate with an increased information dissemination in the network (i.e., structural connectivity does not subserve functional connectivity). This distinction between potential influencers (well-connected members) and effective influencers on StockTwits is attributed, in our examinations, to the presence of a unique content curation algorithm, called the “Watchlist stream.” This content algorithm enables members to capitalize on the network’s momentum through the straightforward act of tracking stocks rather than individuals. Nevertheless, StockTwits members who obtain official account status consistently attract substantial followings and effectively drive information flow within the network. This indicates that official account status is a key determinant of influence.

When we study to what extent the nature of the message may enhance influence, we find that message content and sentiment appear to matter for influence as much as location and status. Specifically, we assess whether messages cite news published by established partisan media outlets (i.e., left-center and right-center). This allows us to determine to what extent perceived unbiasedness (as a proxy for authority) of the medium that a message refers to has an impact on information spread in the network. We also look at the sentiment of the message: whether it concerns Bullish, Bearish or neutral news. The conclusion to emerge from this analysis is that referring in one’s message to a story from a partisan news outlet significantly boosts the likelihood of the message becoming viral in the network. Labelling a message as “Bearish” also increases its likelihood of garnering influence, regardless of whether one is a star in the network. Interestingly, our explorations indicate that very poorly connected members significantly exploit negative sentiment and partisan nature to enhance their influence, as opposed to well-connected members (potential influencers).

When we study the presence and performance of investment advisors in the network, we find that over 87% of investment advisors who become official accounts successfully transform into potential influencers, amassing a considerable following within the network. Interestingly, almost one-quarter of the entire network population comprises followers of professionals, highlighting their crucial role in shaping the network structure. The StockTwits community, devoid of explicit incentives for well-connected users, appears to acknowledge the contributions of investment advisors. The likelihood of a message from a well-connected investment advisor going viral in the network is over 469% higher than that from a regular member. This likelihood even increases by 366% if an investment advisor is registered with either the SEC or state securities administrators (RIA). These findings are not only sizeable, but also highly significant. Notably, we also find that professionals' influence *does* not stem from the partisan nature or sentiment of their messages.

Despite professionals' success in gaining influence, we find that investment advisors—particularly RIAs—have substantially reduced their posting activity since 2018. This trend is also reflected in their sharp exit and fewer new accounts, even as the platform expanded from 1.2 million users in 2017 to 3 million in 2020. Our analysis indicates that this decline stems primarily from increased SEC and FINRA regulatory pressures.

To disentangle the impact of external regulatory constraints on the disintermediation of professional advice from potential changes in the platform or user behavior, we conduct a series of Difference-in-Differences (DiD) analyses—encompassing both standard and dynamic two-way fixed effects (TWFE) specifications with time and user fixed effects. The SEC policy event is dated July 2018, when the Commission reaffirmed amendments to federal securities laws governing professionals' use of social media and announced five related enforcement settlements. The treatment group includes RIAs and unregistered investment advisors (UnRIAs), with all others as the control group. Our analysis focuses on the effect of the SEC policy on professionals' posting activity, over the period from July 2018 to December 2020.

We find that prior to the regulatory intervention, RIAs and UnRIAs posted 487% and 678% more messages than regular members, respectively. Following the SEC restrictions, however, their activity declined significantly, with RIAs posting 53% fewer messages and UnRIAs 42% fewer than other members. A series of robustness checks using DiD analyses with inverse propensity weighting (IPW) and common-support trimming further confirm the chilling effect of regulatory scrutiny.

As such, regulatory enforcement not only reduced advisors' direct engagement on StockTwits but also induced a chilling effect on their participation. This effect is evident because (a) the decline occurred despite RIAs' sustained content production and network influence during the four years preceding the SEC action, reflecting their prior efforts to expand business through social media, and (b) regulators explicitly stated their intent to interpret rules flexibly to facilitate advisors' social media communication with investors (FINRA, 2010). Advisors' overestimation of SEC enforcement risks likely contributed to this outcome, consistent with salience bias (Bordalo, Gennaioli and Shleifer, 2013) and local thinking (Gennaioli and Shleifer, 2010).[2]

Notably, we find that professionals, unlike other users, avoid partisan or emotionally charged content. Combined with prior evidence that their posts more accurately predict next-day stock returns than those of other StockTwits users (Cookson *et al.*, 2024), this suggests that their withdrawal likely reduced the informativeness of socially sourced investment advice, disadvantaging retail investors. However, conflicts of interest and poor trading performance are well documented among advisors (Christoffersen, Evans and Musto, 2013; Linnainmaa, Melzer and Previtero, 2021). As Chalmers and Reuter (2020) argue, whether conflicted advice is better than none depends on institutional context. Our findings, thus, call for external validation, and point to a promising direction for future research.

Our paper contributes to several streams of literature. First, it contributes to the literature on investment advisors' practices in the era of social media, automation, and AI. For example, Greig *et al.* (2024) show that human advisors in hybrid robo-advisors continue to add unique

[2] Under salience bias, advisors overweight the perceived costs of regulatory enforcement relative to the benefits of social media use. Under local thinking, they draw on prior SEC enforcement experiences—even unrelated to social media—when judging the risks of using social media after regulatory actions.

value. Duguay *et al.* (2024) show that the staggered launch of Twitter accounts by the SEC's regional offices deterred advisor misconduct. Cao *et al.* (2024) show that AI cannot replace stock analysts, as humans excel in nuanced judgments. This paper is the first to examine whether investment advisors shape social media to expand their business and whether this activity is affected by SEC regulatory intervention.

Second, our paper contributes to the literature on the diffusion of investment advice among retail investors through financial social networks. For example, Hirshleifer, Peng and Wang (2025) test the social churning hypothesis on StockTwits, showing that earnings announcements initially discussed by well-connected users generate greater reply activity and disagreement within the network. Cookson, Engelberg and Mullins (2023) find that bullish StockTwits users form "echo chambers" by following like-minded peers. This paper is the first to examine virality and the determinants of influence in investment-focused social networks. We show that, unlike on other social media platforms, a higher follower count does not necessarily translate into greater information dissemination.

Finally, our study builds on the literature examining the rise of a new generation of retail investors and their evolving investment preferences. These investors trade more frequently, often driven by attention (Barber *et al.*, 2022; Eaton *et al.*, 2022; Welch, 2022), increasingly use options with poor outcomes (Bryzgalova, Pavlova and Sikorskaya, 2023), and heavily rely on social media for trading decisions, where correlated attention predicts negative returns (Bali *et al.*, 2021; Bradley *et al.*, 2024; Cookson *et al.*, 2024; Chen, Peng and Zhou, 2025). Our paper extends this literature by showing that, despite the new generation of retail investors' reliance on peer insights via social media, they still place significant value on professional advisors' input—and that regulatory enforcement has contributed to the disintermediation of such advice.

The paper proceeds as follows. Section 2 explains StockTwits and the raw data; Section 3 analyzes determinants of influence; Section 4 examines investment advisors' social influence; Section 5 assesses regulatory impacts on their social media use; and Section 6 concludes.

2 The Data

2.1 StockTwits

StockTwits is the second largest finance-oriented social media, boasting over 6 million registered members in 2023.[3] Notably, StockTwits stands out as the most sophisticated social finance community, emphasizing transparency in users' profiles and professional investment strategies. Originally designed for microblogging with a 140-character limit in 2008, the character limit was later expanded to 1000 characters in 2019. StockTwits users share real-time ideas and news stories for stocks and cryptocurrencies using "cashtags" (e.g., $AAPL and $XRP.x for Apple Inc. and Ripple), a suggestive combination of "cash," and the "hashtag" popularized by X (formerly known as Twitter). Until December 2020, more than 1,200 securities were listed on the platform, each with a minimum of 2,000 posts. Additionally, users may attach predefined "sentiment labels" to their posts, such as Bullish and Bearish, providing an explicit indicator of their views and aiding others in interpreting the message.

In the StockTwits network, users can selectively follow designated individuals, creating a stream of messages from followees to followers known as the *People stream*. The decision to follow someone provides network *structure*. Additionally, users can receive tweets only topics-wise (i.e., tweets about a stock/cryptocurrency) through the *Watchlist stream*. For example, users interested in Apple Inc. receive all messages related to it, irrespective of any followee-follower relationships between the authors (originators) and recipients. The resulting transmission of information is therefore based, not on network structure, but on a mapping from topics to subscribers. We refer to this as *bottom-up content curation*.

StockTwits also offers users *top-down content curation*, designed to facilitate user experience, generating streams such as *Suggested*, *Charts*, and *Trending*. Messages are passed,

[3] WSB is the largest social finance with over 14 million registered members in 2023.

not based on network structure or topic, but on what StockTwits deems valuable to increase network activity.[4]

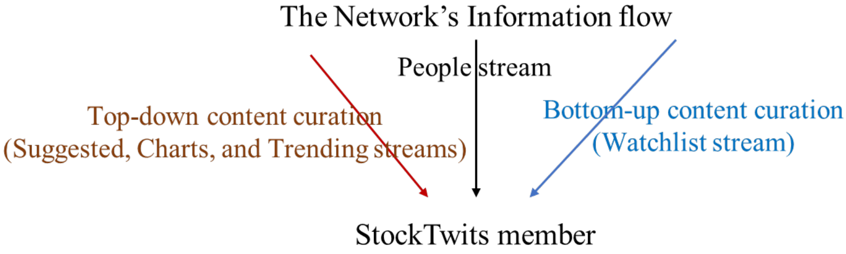

Fig. 1: Information Flow

A StockTwits member who engaged (liked, replied, or reshared) with a message must have received it through either the People stream, bottom-up content curation, or top-down content curation algorithms. As standard in literature, we consider a tweet to be received via the People stream if the member follows the message's originator (Goel, Watts and Goldstein, 2012; Goel *et al.*, 2016; Vosoughi, Roy and Aral, 2018). In the absence of a followee-follower relationship, we attribute the message origin to the Watchlist stream (bottom-up curation) if the member follows stocks tagged in the tweet. Otherwise, top-down curation is considered as the origin of the tweet.

2.2 Our StockTwits Data

We restrict the sample to messages pertaining to a random selection of 129 U.S. stocks over the period January 2014 to December 2020. The list of sampled firms is provided in Table A. 1 of the Appendix. To mitigate selection bias and enhance representativeness, we implement a three-stage stratified sampling procedure, stratifying firms by industry classification and firm size. In the first stage, firms are grouped into 11 sectors based on the Global Industry Classification Standard (GICS).[5]

Second, we labeled all firms of each sector with respect to size as of the end of 2019, using three predefined labels, namely, "large," "medium," and "small."[6]

[4] The content curations algorithms of the network, use "Stocktwits ranking algorithms" to shuffle tweets into what it proposes is a better order i.e., "top Tweets." For example, "Suggested" shows the top tweets posted by the most popular and influential Stocktwits members. "Charts" presents the top tweets which include charts and pictures. "Trending" helps users to discover trending tickers in the network, which are undergoing a significant Social Attention through message volume, likes, replies, and retweets.
[5] GICS defines 11 sectors including Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, Information Technology, Real Estate, Communication Services, and Utilities Sector.
[6] We used Nasdaq's definitions of market capitalization (Market Cap) classes for firms. Where, a firm is defined large if Market Cap > \$10B, medium if \$2B ≤ Market Cap ≤ \$10B, and small if Market Cap < \$2B.

Table 1: Characteristics of StockTwits Data

Table reports summary statistics for our StockTwits data. Panel A shows dimensions of the data. Panel B presents summary information on coverage by stock. Panels C and D display summary information for "regular users" and "official users," respectively.

Panel A. Dimension of the data

	Number of Messages			
firms	Sentiment			Total
	None	Bearish	Bullish	
DOW Jones	2,180,386	278,838	908,740	3,367,964
Large	1,696,307	360,236	922,261	2,978,804
Medium	304,016	62,961	203,895	570,872
Small	1,143,972	110,964	1,100,352	2,355,288
Total	5,324,681	812,999	3,135,248	9,272,928

Users	Number	%Total
Regular	233,475	99.8%
Official	492	0.2%
Total	233,967	100%

Days	
Trading	1,129
Non-Trading	484
Total	1,613

firms	**Number**	**%Total**
DOW Jones	30	23.3%
Large	33	25.6%
Medium	33	25.6%
Small	33	25.6%
Total	129	100%

Panel B. Message coverage frequency per firm

	Obs.	Mean	SD	Min	25%	Med	75%	Max
Number of messages per day	115,052	81	810	1	3	12	44	109,778
Number of messages	129	23,961	50,352	15	266	3,362	22,828	506,156
Number of Bearish messages	129	6,302	15,269	1	67	471	6,224	112,216
Number of Bullish messages	129	24,304	45,397	22	225	1,997	22,828	268,332
Number of messages w/o label	129	41,277	68,812	543	2,757	7,673	59,237	506,156

Panel C. Regular users

Characteristic	Mean	SD	Min	25%	Med	75%	Max
Number of messages	37	434	1	1	4	16	125,862
Number of followers	80	1,505	0	1	3	11	195,308
Number of followees	34	102	0	2	11	58	9,981
Number of securities in Watchlist	38	55	0	6	18	44	308
Number of likes	707	3,507	0	13	78	385	810,222
Total days active	537	685	10	57	231	796	4,162

Panel D. Official users

Characteristic	Mean	SD	Min	25%	Med	75%	Max
Number of messages	1,302	4,070	1	20	112	556	45,749
Number of followers	37,373	48,100	7	5,146	23,000	52,310	495,768
Number of followees	319	1,079	0	27	87	211	10,000
Number of securities in Watchlist	21	44	0	0	7	20	300
Number of likes	1,635	6,354	0	12	82	646	84,497
Total days active	2,168	1,171	10	1,241	2,225	3,204	4,197

Third, we randomly selected 9 firms from each sector, including 3 large, 3 medium, and 3 small firms. This consists of 99 randomly selected firms. To complete the sampling process, we additionally included the 30 firms comprising the Dow Jones Industrial Index (Dow) as of the end of 2019. These firms are incorporated into the stratified sample to ensure adequate representation of companies with potentially high StockTwits coverage, which may not be guaranteed through random selection alone.

Our dataset provides a broad and unbiased representation of StockTwits activity, covering over 60% of the platform's user base and more than 20% of all messages. The sample is also well-distributed across firm size and industry classifications. After removing spam content, the final dataset comprises over 9.2 million messages posted by approximately 233,000 unique users.[7] For each message, the dataset includes the timestamp, self-reported sentiment, referenced stock tickers, and the full text content, including any embedded URLs. It also records the identity of the original poster, as well as the identities of users who replied to, liked, or reshared (retweeted) the message. Table 1 presents the summary statistics of our StockTwits data.

2.3 StockTwits Professionals

To examine the presence and influence of professional investment advisors on StockTwits, we conduct a manual review of official account profiles within our dataset, focusing on verified users to ensure the accuracy of identity and professional affiliation. StockTwits designates official accounts with a gold badge (), enhancing the credibility of profile information. We analyze external links provided in these profiles—such as business websites and LinkedIn pages—and cross-reference them with FINRA's BrokerCheck to verify Registered Investment

[7] StockTwits actively removes spammer accounts, biasing the collected data. However, to decrease the noise of spambots, specifically we restricted on messages of "regular users" who have been active more than 10 days in the network. We did not limit on messages of "official users," as their profiles are confirmed by StockTwits.

Advisor (RIA) status. Fig. 4 illustrates an example of this verification process using official account profiles.

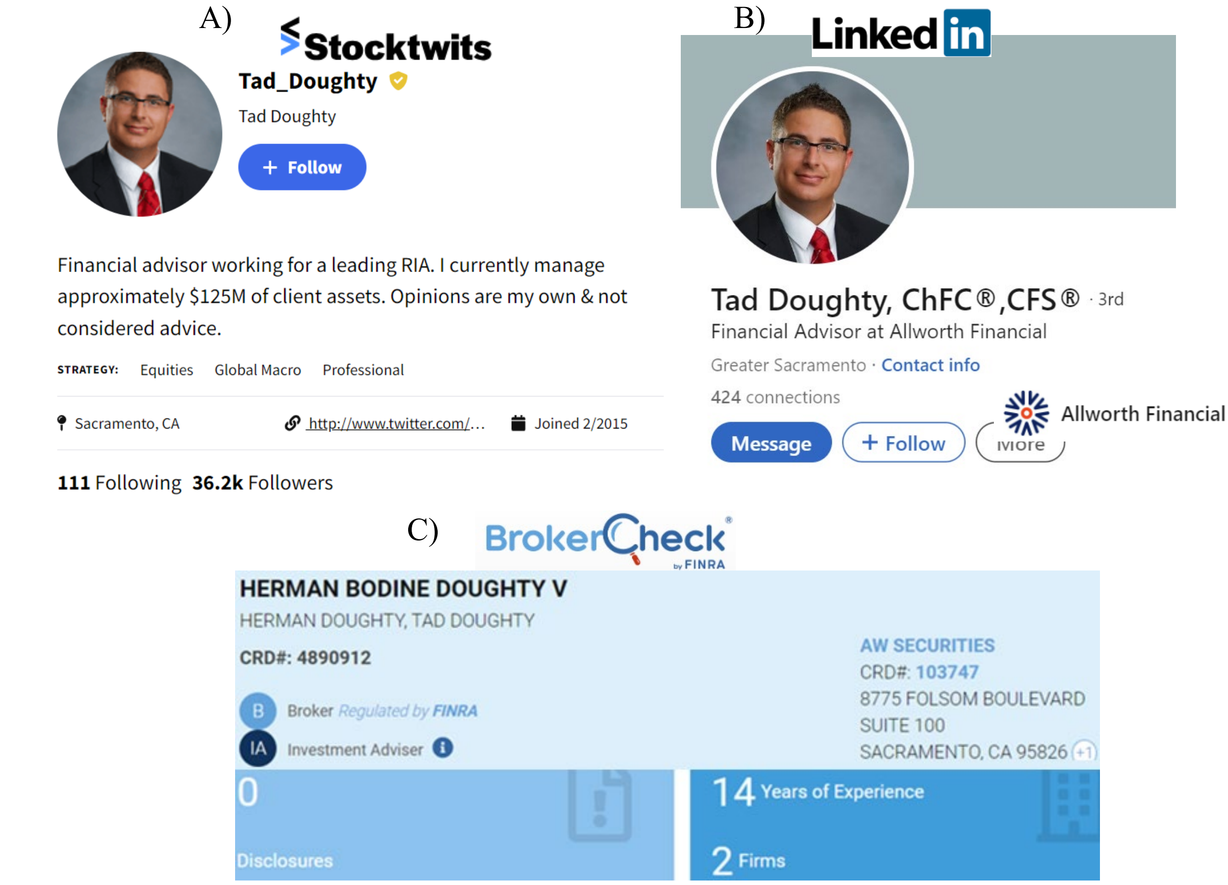

Fig. 2: Examination of Official Account Profiles

Figure depicts our investigation of investment advisors on StockTwits using official account profiles. Panel A displays the StockTwits profile of Tad Doughty, a verified user with 36.2K followers, identifying himself as an RIA based in Sacramento, California. Panel B validates this information through his LinkedIn profile, indicating employment at Allworth Financial, L.P. Panel C presents a screenshot from FINRA BrokerCheck, confirming his active RIA status and 14 years of experience in the investment advisory industry.

Based on our investigation, we classify official accounts into three categories: Registered Investment Advisors (RIAs), Unregistered Investment Advisors (UnRIAs), and other official users. The UnRIA group includes professionals affiliated with investment advisory firms or platforms that offer investment-related tools or content but are not registered with FINRA. Official StockTwits accounts not engaged in professional investment advising are categorized as "other." Table 2 reports summary statistics for the three categories of official users. As shown, we provide detailed metrics for each category, including the number of messages, followers, followees, securities in the watchlist, likes, and the total number of active days.

Table 2: Characteristics of Official Accounts

The table presents summary statistics for official users categorized into three groups: registered investment advisors (Panel A), unregistered investment advisors (Panel B), and other official users (Panel C). Classification into these groups is determined through our manual examination of StockTwits profiles, cross-referenced with information from LinkedIn and FINRA BrokerCheck. The presented statistics are derived from a snapshot of 492 official accounts in our StockTwits data as of December 2020.

Panel A) StockTwits Registered Investment Advisors (RIAs), 75 observations.

Characteristic	Mean	SD	Min	25%	Med	75%	Max
Number of messages	8,243	12,668	3	912	3,524	13,254	87,654
Number of followers	45,653	37,390	11	13,828	33,684	74,357	161,701
Number of followees	260	877	0	26	98	246	7,484
Number of securities in Watchlist	18	42	0	0	5	20	300
Number of likes	1,108	3,073	0	12	66	610	18,340
Total days active	2,948	916	10	2,549	3,141	3,658	4,197

Panel B) StockTwits Unregistered Investment Advisors (UnRIAs), 261 observations.

Characteristic	Mean	SD	Min	25%	Med	75%	Max
Number of messages	17,705	46,937	17	1,267	4,364	13,831	531,122
Number of followers	36,385	39,477	57	9,166	24,602	49,469	309,292
Number of followees	351	1,198	0	22	83	224	10,000
Number of securities in Watchlist	21	45	0	0	6	17	300
Number of likes	1,353	4,646	0	12	76	630	44,181
Total days active	2,857	924	58	2,189	3,025	3,658	4,105

Panel C) Other Official accounts, 156 observations.

Characteristic	Mean	SD	Min	25%	Med	75%	Max
Number of messages	16,293	41,036	3	336	2,136	9,500	272,207
Number of followers	35,045	63,302	7	248	12,102	43,202	495,768
Number of followees	296	955	0	30	85	195	9,999
Number of securities in Watchlist	24	43	0	1	9	27	297
Number of likes	2,362	9,294	0	13	96	701	84,497
Total days active	2,150	876	650	2,140	2,738	3,501	4,105

3 Determinants of Influence on StockTwits

This section investigates the percolation of information and the determinants of influence on StockTwits, addressing our first research question. Section 3.1 analyzes *structural connectivity* to identify potential influencers—highly connected users who, according to network theory, are *likely* to drive extensive information diffusion. Section 3.2 examines *functional connectivity*, capturing the actual flow of messages within the network using a cascade model. This framework enables an assessment of the alignment between structural and functional connectivity in Section 3.3. A lack of alignment would suggest that effective influence on StockTwits does not necessarily correspond to network centrality or follower count. Finally, Section 3.4 explores the determinants of effective influence using a series of random-effects logistic regressions, estimating the probability of a message being retweeted based on characteristics of the original poster and the content of the message.

3.1 Structural Connectivity

Our analysis begins by identifying highly connected users within the network through an examination of *structural connectivity*, which refers to the underlying network topology—specifically, the pattern of who follows whom. Analyzing structural connectivity enables the identification of potential social influencers, as well-connected nodes are theoretically more capable of disseminating information widely. Network topologies are typically classified as either *scale-free* or *random*. In a scale-free network, the distribution of connections per node follows a power-law, implying that a small proportion of nodes account for a disproportionately large share of connections. According to network theory, such well-connected nodes are likely to function as potential influencers, driving the flow of information within a scale-free network (Pastor-Satorras and Vespignani, 2001; Barthélemy *et al.*, 2004; Barabási, 2009; Clauset, Shalizi and Newman, 2009). In contrast, within networks characterized by a random topology, each node has an approximately equal probability of becoming a potential influencer (Watts and Strogatz, 1998; Watts, 2002).

Many real-world networks, such as X (Kwak *et al.*, 2010; Myers *et al.*, 2014; Bild *et al.*, 2015), exhibit key features of *scale-free networks* (see Fig. 1(A); (Pastor-Satorras and Vespignani, 2001; Barthélemy *et al.*, 2004; Barabási, 2009)). This means that the number of connections k of a randomly chosen node follows a power-law distribution (beyond some critical k*, $P(k) \sim k^{-g}$; $g > 0$ and $k \geq k^*$). Intuitively, a relatively small but significant number of nodes are responsible for most connections. The phenomenon gives rise to the so-called law of (Vilfredo) Pareto, who observed that 20% of the Italian population owned 80% of the land.

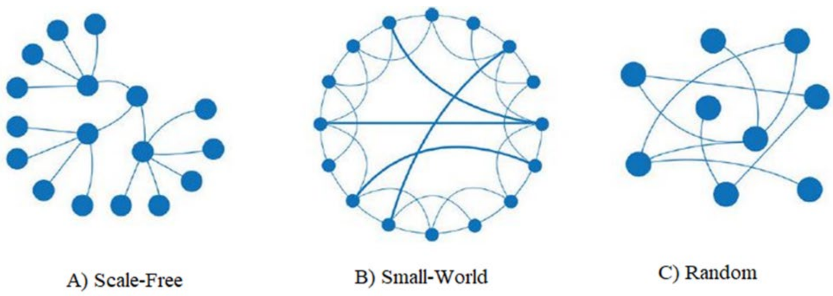

Fig. 3: Network Structures

Scale-free network (A) has high clustering. Random network (C) lacks hierarchy; average and well-connected nodes share influence. Small-world network (B) blends scale-free's clustering with random connections. Source: Needham and Hodler (2019).

Following the standard statistical protocol outlined by Clauset *et al.* (2009) and Broido and Clauset (2019), we show that StockTwits constitutes a scale-free network structure. This suggests that well-connected members are likely to act as potential influencers, facilitating the widespread dissemination of information within the network. Table 3 displays the results of the statistical tests of a pure power law against various alternatives, including the best-fitting one, a truncated power law ($P(k) \sim k^{-g}\, e^{-\lambda k}, k \geq 15, g = 1.77, and\ \lambda = 2.56e - 07$).

As shown in Table 3, the best-fitting distribution generates g = 1.76. Since g < 2, StockTwits is only a weak scale-free network. This implies that the distribution of connections is well-approximated by a power law, except in the far-right tail, where the distribution exhibits an exponential cutoff. Fig. 5 illustrates this pattern: beyond 100,000 followers,[8] the empirical distribution function deviates sharply below the best-fitting power law in log-log space. Nevertheless, the presence of a power-law regime is evident.

[8] For comparison: the average "Official" user has ~37,000 followers (see Table 1).

Table 3: Tests of power law behavior in the follower count distribution

This table represents the results for statistical support for the power law hypothesis (P-Law), "scale-free structure," for the follower count distribution on StockTwits. The goodness-of-fit for the power law ($P(k) \sim k^{-g}, k \geq 15,\ g = 1.76$) is compared with other alternative distributions fitted to the same part of the far tail using the likelihood-ratio test. The p-value for the significance of the likelihood ratio is also shown in the table. Statistically significant p-values are denoted in **bold**. Positive log-likelihood ratios (and greater than 1) indicate that the power law model is favored over the alternative. The final column of the table represents our judgement of the statistical support for the power law hypothesis for the empirical data.

Test (H_0 vs H_1)	Log-Likelihood Ratio	p-value	Supported Hypothesis
P-Law	-	**0.096**	H0
P-Law vs Log-normal	1.485	**0.000**	H0
P-Law vs Log-normal (positive)	1,425	**0.000**	H0
P-Law vs Exponential	114,476	**0.000**	H0
P-Law vs Truncated P-Law	-2.968	**0.015**	H1 (Estimates: $P(k) \sim k^{-g}\ e^{-\lambda k}, k \geq 15$ $g = 1.77, and\ \lambda = 2.56e - 07$)

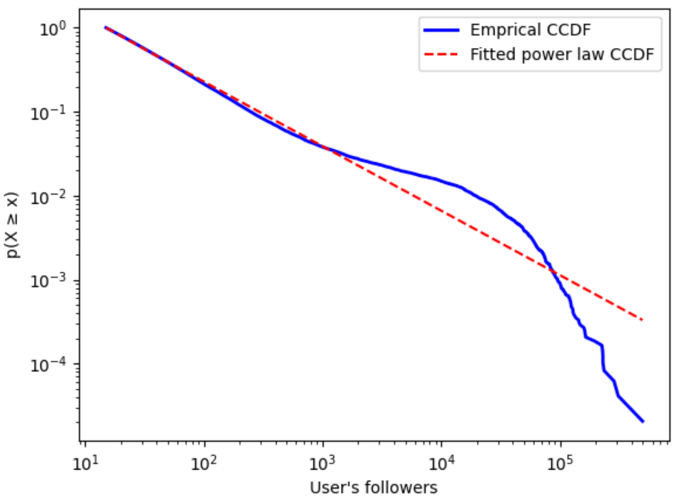

Fig. 4: Empirical CCDF of Follower Count Data and Fitted Power Law

This figure shows the complementary cumulative distribution function ($p(X \geq x)$), of follower count (X) in StockTwits (blue), as well as the best-fitting power law distribution (red).

Comparing message volume by user group, we find that the top 10% of users—those with at least 1,000 followers—contributed approximately 1,000,000 messages, whereas the bottom 90% generated only about 600,000 messages. Rather than observing the classic 80/20 Pareto distribution, the data reflect a "60/40" pattern, indicating a more moderate concentration of content generation; see Table 4.

Hereafter, we refer to users with at least 1,000 followers as "potential influencers." This threshold is selected because, as is clear from Fig. 5, this is where the tail of the distribution of follower count becomes heaviest; above 10,000, the tail reverts to becoming exponential. According to network theory, users exceeding the 1,000-follower threshold are expected to generate a disproportionate share of messages and are therefore likely to be among the most effective influencers in the network.

In theory, this is favorable for the investment advice industry, as the majority (over 75%) of professionals have successfully accumulated more than 9,000 followers within the network; see Table 2 (A and B). However, two key points merit consideration: Firstly, the efficacy of social influencers depends on the broad dissemination of their messages within the network, a factor not necessarily linked to follower count (functional connectivity may not subserve structural connectivity). Secondly, StockTwits exhibits a weak scale-free structure, which may invalidate network theory predictions. Thus, it remains to be proven whether well-connected members of the network are also effective influencers.

It may also be important to distinguish between the potential influencers and the official users. Because the latter too have a significant number of followers (on average, over 37,000; see Table 1 (D)), they are mostly a subset of the former. Later, we shall determine to what extent there is a difference between a professional investment advisor and a non-professional well-connected node.

3.2 Functional Connectivity

In this section, we apply network science to measure the actual flow of information (functional connectivity) on StockTwits using the *cascade model*. This approach is crucial because comparing structural and functional connectivity enables us to determine whether potential influencers (well-connected members) are also effective influencers. To assess functional connectivity, we construct the cascade for every message (tweet) in our dataset. A cascade begins when a user posts a tweet, referred to as the *original poster* (or original author). Subsequent users expand the cascade by resharing (retweeting) the original tweet. Analyzing these cascades allows us to identify *effective influencers*, from whom the majority of large cascades originate.

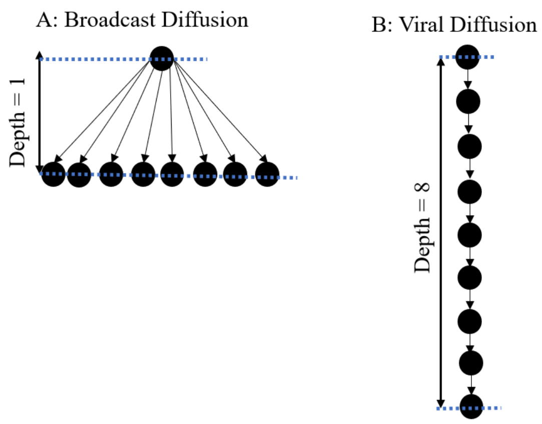

Fig. 5: Geometry of Cascades

Cascade geometry spans the extremes of broadcast and viral dynamics. In the broadcast model (A), a single node initiates a burst of adoptions, with the originator contributing the majority of the total adoptions. In contrast, the viral model (B) involves person-to-person diffusion, creating a branching structure where each individual plays a role in the overall adoption process.

The motif ("geometric structure") of a cascade can be identified through *size*, *depth*, and *structural virality*. The *size* of a cascade shows the number of unique users involved in the cascade, *including the originator.* A cascade can be as small as size one, in which case no one re-tweeted. Fig. 6 depicts a cascade in StockTwits. The size of the cascade equals 10 (remember that the size includes the original tweeter in the count).

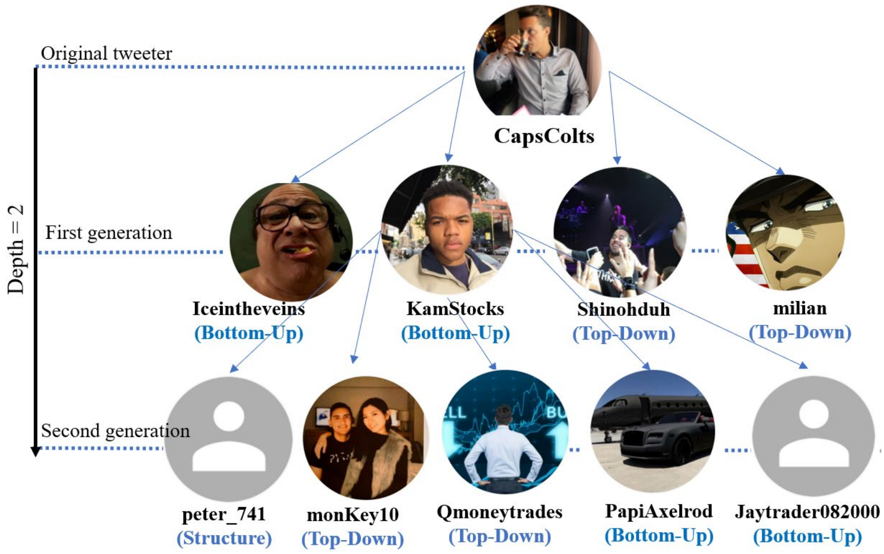

Fig. 6: Example of A Cascade in the StockTwits Network

The figure illustrates the tree structure of a cascade on StockTwits. In this case, the original poster (CapsClots) tweeted about Remark Media Inc. ($MARK), linking to a news article published by CNN.com on June 5, 2020. The cascade includes 9 adopters who further expanded the cascade through retweeting, resulting in a total cascade size of 10. Each adopter received the post either through content curation (both top-down and bottom-up) or because of their structural connection via the People stream (Structure).

The *depth* of a cascade measures the number of *generations* in the corresponding cascade tree. To determine the depth, one starts from each final recipient (i.e., the last users to retweet) and counts the number of retweets required to trace back to the originator of the message. The depth is then defined as the maximum of these counts across all final recipients. The cascade in Fig. 6 spans 2 generations of adopters (the depth of the cascade is 2). The first generation of adopters reshared the post directly from the original poster, while the second generation of adopters consists of users who retweeted the original post from the first-generation adopters.

Structural virality, as proposed by Goel *et al.* (2016), summarizes the geometric structure ("motif") of a cascade. The geometry of a cascade varies along two conceptual extremes: broadcast and viral. In a broadcast cascade, popularity is driven by a large burst of adoptions from a single parent node (see Fig. 2(A)), with the primary contribution to total adoption coming from the original poster. In contrast, a viral cascade exhibits multigenerational branching, where multiple individuals contribute to the total adoption, each responsible for a portion of the overall spread (see Fig. 2(B)). Goel *et al.* (2016) define the structural virality as the average distance between all pairs of nodes in a diffusion tree. For a cascade with $n > 1$ nodes (the message is retweeted at least once), structural virality (v) is defined as follows:

$$v = \frac{1}{n\,(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}, \qquad (1)$$

Where d_{ij} stands for the length of the shortest path between nodes i and j. Higher values of structural virality indicate that adopters are, on average, farther apart in a cascade and the geometry looks like that of viral diffusion (Fig. 2(B)). Broadly speaking, $v \approx 2$ corresponds to the broadcast diffusion, and $v > 2$ corresponds to viral diffusion. The structural viralities equal 1.78, 3.33, and 2.13 in Fig. 2(A), Fig. 2(B), and Fig. 6, respectively.[9]

To construct cascades with StockTwits data, we proceeded as follows. First, we defined an adopter as either the original author of the content (root node) or a "child" of another adopter in the tree structure. An adopter is someone who (re-)tweeted a message. With respect to the "children," we shall refer to the parent as "inferred parent." All children of a particular inferred parent are identifiable in two types: (i) those who officially retweeted the parent's post, and (ii) those who did not retweet that parent's message but for whom there is evidence suggesting he/she has learned the content from the parent. Attribution of children of type (i) is straightforward because StockTwits records all the users who officially retweeted a post. However, type-(ii) children are more difficult to identify, as StockTwits classifies these posts under originally authored content. Here, we consider an adopter as type-(ii) child of a particular inferred parent if the adopter posted the content later in time than the parent, mentioning the

[9] Fig 2(A): (8*1+8*(1+7*2))/(9*8) = 1.78; Fig 2 (B): ((1+2+3+4+5+6+7+8)+(1+2+3+4+5+6+7+1)+(1+2+3+4+5+6+1+2)+(1+2+3+4+5+1+2+3)+(1+2+3+4+1+2+3+4)+(1+2+3+1+2+3+4+5)+(1+2+1+2+3+4+5+6)+(1+1+2+3+4+5+6+7)+(1+2+3+4+5+6+7+8))/(9*8)= 3.33; Fig. 6: ((4*1+5*2)+3*(1+3*2+5*3)+(1+3*2+5*1)+5*(1+5*2+3*3))/(10*9)= 2.13.

parent's username in the post using the @ sign. This identification of children is repeated from root to leaf for each adopter.

Moreover, as standard in literature, we focus our study on cascades where the original tweet refers to websites via URL links. This is because the literature considers a tweet to be informative if it contains a hyperlink to a website (Goel *et al.*, 2012; Goel *et al.*, 2016; Vosoughi *et al.*, 2018). A unique piece of content (message pointing to a website) can generate multiple cascades. The number of diffusion cascades indicates the number of times the content is independently tweeted by new root node adopters.[10] Specifically, whenever we observe an adopter who posted a message, neither through retweeting nor mentioning other adopters, we consider that adopter as a "root node" for a new cascade in the network.

We recorded 1,585,082 cascades, spread by about 61,000 unique originators. Indeed, comparing StockTwits with X (formerly Twitter) is interesting, as both use microblogging with similar structures and content algorithms. However, X supports broad social interactions, while StockTwits focuses solely on investment advice .As such, we adopt the relevant statistics for X directly from Goel *et al.* (2012).

The left panels in Fig. 7 display the proportion of messages that were retweeted on StockTwits (Panel A) and X (Panel B). Re-tweeted messages constitute a small share of total posts, with their frequency on X (7%) being more than twice that on StockTwits (3%). The right panels in Fig. 7 present the distribution of *adoption size* among re-tweeted messages, defined by the number of distinct users who further shared the content. Specifically, 87% of re-tweeted messages on StockTwits are shared only once (cascade size = 2), compared to 71% on X. This is followed by two re-tweets (size = 3; 9%) and three re-tweets (size = 4; 2%) on StockTwits.

[10] For instance, the message that led to the CapsClots cascade depicted in Fig. 6 led to two additional cascades: we identified two more users who independently tweeted the CNN article for $MARK, without discernible link to the CapsClots cascade.

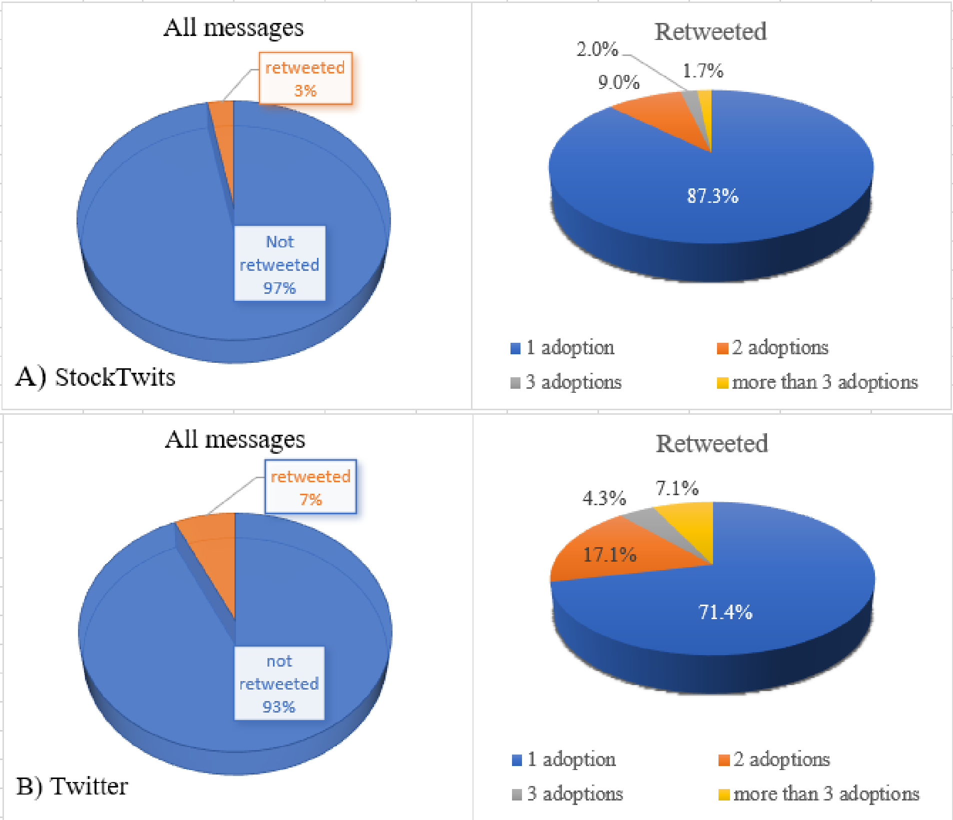

Fig. 7: Message Diffusion on StockTwits vs. X (formerly Twitter)

The left panels show the proportion of messages that were retweeted on StockTwits (Panel A) and X (Panel B). Retweeted messages represent a small fraction of total posts—3% on StockTwits and 7% on X. The right panels display the distribution of adoption size among retweeted messages, measured by the number of distinct users who further shared the content. The adoption statistics for X are taken directly from Goel *et al.* (2012).

As such, the distribution of cascades reveals some similarities between StockTwits and X. In both networks, most messages are never re-tweeted, and among those that are, the majority are re-tweeted only once. However, X exhibits stronger structural connectivity than StockTwits. The similarity in cascade patterns may reflect the more prominent role of content curation streams in social finance platforms like StockTwits, compared to general-purpose networks such as X. As shown in the subsequent analysis, this appears to be the case.

3.3 Alignment of Structural and Functional Connectivity

After measuring both structural and functional connectivity on StockTwits, we next examine the extent to which the network's structure supports functional connectivity. Using our cascade data, we trace the source of each message received by an adopter; see Fig. 3. We distinguish between adoptions that occur through the network's structural connections (i.e., the People stream) and those that occur outside the structural network, including bottom-up content curation (the Watchlist stream) and top-down content curation algorithms (the Suggested, Charts, and Trending streams).

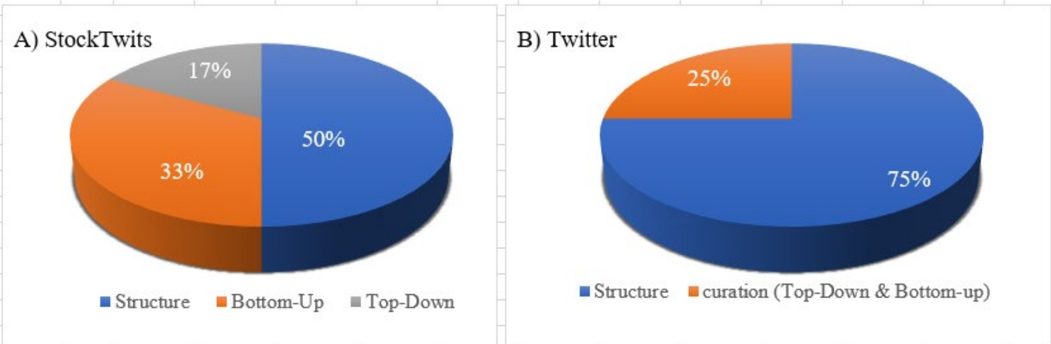

Fig. 8: Adoption Frequency, Through Network Structure or Through Content Curation

Fig. 8(A) and Fig. 8(B) show the percentage of adoptions through structure and curation (the top-down and bottom-up), in StockTwits and X (formerly Twitter), respectively.

As shown in Fig. 8(A), we find that somewhat less than 50% of messages are adopted from the structure: adopters (children) are following inferred parents. Among the remaining adoptions, 2/3 originate topic-wise (adoptions due to stocks tagged in messages), via the Watchlist stream (bottom-up content curation). As such, message flow through the network caused by the bottom-up content curation is sizeable, reducing the relevance of the network structure in social finance. The remaining adoptions—at most one-sixth of the total ((1/2) × (1/3))—can be attributed to top-down content curation. This suggests that the influence of StockTwits' top-down content curation (People Stream) on network activity is relatively limited.

Let us compare this to X (Fig. 8(B)): three quarters of adoptions arrive through the structure of the network, while the remaining one quarter arrives through the top-down curation; see the appendix in the study by Goel *et al.* (2016). Compared to StockTwits, structure in the X network is sizeably more important for functional connectivity. As a result, in StockTwits, one does not need to be well-connected to potentially experience influence; instead, one can leverage the Watchlist stream to reach a broad audience beyond their direct followers. For example, by tagging $AAPL in a post, a user can disseminate the message to over 600,000 followers of AAPL on StockTwits through the Watchlist stream. This then provides an explanation for why StockTwits is a weak scale-free network. Conversely, the occurrence of adoptions through the curation channels suggests that well-connected nodes within the network might not wield the expected influence. As a preliminary step, we examine the distribution of cascades with respect to the structural connectivity of their originators. Table 4 presents the message frequency from potential influencers (well-connected users) and other network members, along with the corresponding resharing frequency for these messages.

Table 4: Potential Influencers vs Other Members

User Type	Observation	Frequency that messages are reshared by others
Potential influencers (#followers ≥ 1000)	977,961	2%
Non-central Users (#followers < 1000)	607,121	4%

As shown in the table, about 4% of non-central nodes are re-tweeted, while only 2% of messages from potential influencers produce cascades of size larger than 1. Since messages can cascade beyond network structure, we need to study in more detail what creates effective influence, over and beyond a count of one's followers.

3.4 What Determines Effective Influence?

Since StockTwits was found to be a scale-free network, network theory would predict that the potential influencers will also be effective influencers, because of the sheer volume of followers they are connected to. However, our preliminary findings in Section 3.3 indicate that StockTwits exhibits characteristics of a weak scale-free network, where structural connectivity does not support functional connectivity to the same extent as observed in scale-free networks such as X. Our analysis attributes this phenomenon to the presence of the Watchlist stream, indicating that StockTwits users exhibit a pronounced inclination to leverage the network's momentum by primarily monitoring stocks rather than individuals.

The disconnect between structural and functional connectivity could imply that *effective influencers* on StockTwits may differ from *potential influencers*. To examine what determines the influence on StockTwits, we run a logistic regression. We allow for random intercepts (random effects), thereby controlling for heterogeneity of mean effects at the level of the firms mentioned in the posts.[11] The regression also includes poster (original poster) and year (time) fixed effects.

The dependent variable is whether the original message of a cascade was re-tweeted (yes = 1; no = 0). As explanatory variables, we included:

- A dummy for *Potential influencer* (i.e., users with more than 1,000 followers)
- A dummy for *Official users* (i.e., nodes of *Official Account* holders, mostly professionals, as well as individuals recognized to be stars by StockTwits). Examples of official Account holders are displayed in Fig. A. 1, Appendix. One would expect the status of the official account to provide additional effectiveness (most official users are potential influencers). An argument against the latter is the concern that professionals are generally viewed with mistrust, as opposed to peers (as observed in the case of WSB during the GameStop short squeeze in early 2021).
- Other explanatory variables related to the content of the message, such as the *Partisan nature* and the *self-declared sentiment label* of the post; both treated as dummy variables.

It is important to investigate the role of message content in generating influence within the network. In particular, we assess the extent to which influence may be shaped by the strategic use of sentiment labels and external links. As a first step, we analyze the sentiment associated with each message. Message sentiment refers to the label that a StockTwits user may assign to

[11] **STATA commands:** xtlogit dependent_variable [independent_ variables], re or" and "margins, over (independent variables) atmeans expression(exp(xb())) noatlegend.

a post, indicating whether the content is Bullish (*Positive sentiment*) or Bearish (*Negative sentiment*) with respect to the stock discussed. Messages without a sentiment label are classified as *neutral*.

We further examine message content by assessing the influence of political bias on information dissemination. To this end, we construct a binary variable, *Partisan Nature*, which equals 1 if a message includes hyperlinks to news articles published by partisan media outlets, and 0 otherwise. *Partisan nature* is defined according to the widely recognized political leanings of the news sources. Media outlets are classified using assessments from three independent fact-checking organizations: Ad Fontes Media, Media Bias Fact Check, and AllSides. Based on these classifications during the period of our study (2014–2020), CNN, MSNBC, *The New York Times* (NY Times), and *The Washington Post* (The Post) are labeled as left-biased, while *The New York Post* (NY Post), Fox News, and *The Wall Street Journal* (WSJ) are identified as right-biased. Messages that cite other sources are classified as politically neutral, or at least less overtly partisan in nature. We ensure that no news outlets cited in our dataset are rated as more politically biased than the seven outlets identified above. Within our data, 9,799 originator messages cite left-biased sources, and 8,488 cite right-biased sources, causing 20,417 cascades (see Fig. A. 2 in the Appendix for the cascade count across main news outlets in our data).[12]

The results of the logistic regression analysis are reported in Table 5. As shown in Table 5(A), the most striking result we obtain is that the odds ratio associated with potential influence is significantly and substantially below one. Specifically, messages originating from well-connected nodes are notably about 20% less likely to be re-tweeted compared to those from less connected nodes. Even if potential influencers sent many more messages (~1,000,000; see Table 4) than non-central users (~600,000), their effectiveness is much reduced. The poorer

[12] Note that a unique news story may trigger multiple independent cascades, leading to a higher cascade count than the number of news stories disseminated on the network.

effectiveness almost compensates for greater activity. About 4% of non-central users are re-tweeted, while only 2% of messages from potential influencers produce cascades of size larger than 1; see Table 4.

In contrast, official users exhibit a significantly and substantially greater impact, with their messages being approximately 75% more likely to be re-tweeted than those of regular users. This finding offers a positive signal for the investment advice industry.

given social finance presents a potential challenge to professional financial advisors, as it may be perceived as a substitute for traditional advisory services. Yet, by actively participating in a social network like StockTwits and obtaining Official Account status, professionals can retain some control over the narrative. However, as shown in Table 5(B), such influence materializes primarily when an Official Account is also well-connected—that is, when the user qualifies as a potential influencer. Specifically, posts from official users with fewer than 1,000 followers (non-central users) have only a 0.5% probability of being reshared. This probability more than doubles, to 1.2%, when the official user also has at least 1,000 followers.

Regarding message content, we conjecture that unbiasedness would be a major determinant of whether a message would be re-tweeted.[13] As shown in Table 5(A), contrary to our expectation, referencing a story from a partisan news outlet significantly raises the likelihood of message retweets. The probability of a message being retweeted increases by ~50% when it refers to a story from a biased (left or right) news source. More importantly, *Negative sentiment* has an even more pronounced effect on influence: the likelihood that a message is re-tweeted increases by ~75% when the originator tags it as Bearish. This is not only sizeable, but also highly significant ($p < 0.001$).

[13] We also studied perceived reliability of the news media using the same media fact checking organizations. However, perceived reliability had no influence on message re-tweeting.

Table 5:The Likelihood of Re-Tweeting

This table shows the results of a random-effects logistic panel regression (panel variables: firms mentioned in messages) estimating the likelihood of retweeting of a post linking to a website on StockTwits, as a function of variables including the characteristics of original tweeter, partisan nature of content, and self-declared sentiment label of the post. Model specification is as follows (**Number of observations = 1,585,082**, Wald χ^2 =- 215,957, with p < 0.001):

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta_1\,Potential\,influencer + \beta_2\,Official\,user + \beta_3\,Negative\,sentiment + \beta_4\,Partisan\,nature + \beta_5\,Partisan\,nature * Negative\,sentiment + \beta_6\,Potential\,influencer * Official\,user + Poster + Year + \varepsilon_0\,,$$

where p is the likelihood of resharing a unique piece of information in the network. α_1 and α_2 indicate random effects per message sender and per firm mentioned in the message, respectively. Original tweeter characteristics are measured through "Potential influencer" and "Official user." "Potential influencer" is a dummy variable equal to 1 if the original tweeter does have more than 1000 followers. "Official user" is a dummy variable equal to 1 if the original tweeter holds an "Official Account." "Negative sentiment" is a dummy variable equal to 1 if the original tweeter has labeled his/her post as "Bearish." "Partisan nature" is a dummy variable equal to 1 if the original tweeter's post linking to a news article published by the media deemed partisan (WSJ, NY Times, The Post, NY Post, CNN, FoxNews, and MSNBC). The regression includes poster and year fixed effects. The odds ratios for β_1 to β_4 are reported in Panel A. "Probability of resharing (prob.)" of a post is based on marginal effects of different combinations of user type as well as nature and sentiment and is shown in Panel B. A likelihood ratio test comparing this model to a standard logit yields **χ2(2)= 16,925** (p<0.0001), justifying the panel structure. The estimated variance components show that **var(constant)= 1.941**, indicating unobserved firm-level heterogeneity in sentiment outcomes.

Panel A) Odds ratios

	Odds Ratio	Std. Err.	Z	p-value	95% Confid. Interval	
Potential Influencer	0.798	0.008	-20.4	<0.001	0.781	0.816
Official user	1.750	0.022	44.7	<0.001	1.707	1.794
Partisan nature	1.499	0.053	11.4	<0.001	1.399	1.607
Negative sentiment	1.748	0.036	27.4	<0.001	1.680	1.819
Constant	0.009	0.001	-36.9	<0.001	0.006	0.011
Var(Constant)	1.941	0.282			1.460	2.582

Panel B) Probabilities related to interaction effects

Interaction	Prob. (%)	Delta-method Std. Err. (%)	Z	p-value	95% Confid. Interval	
Non-centralUser * non-Partisan * not Negative	0.897	0.115	7.77	<0.001	0.671	1.124
Non-centralUser * non-Partisan * Negative	0.535	0.09	5.95	<0.001	0.359	0.711
Non-centralUser * Partisan * not Negative	0.681	0.088	7.77	<0.001	0.510	0.853
Non-centralUser * Partisan * Negative	1.226	0.158	7.76	<0.001	0.916	1.535
Potential Influencer *non-Partisan*not Negative	0.770	0.099	7.76	<0.001	0.576	0.965
Potential Influencer * non-Partisan * Negative	1.375	0.179	7.66	<0.001	1.023	1.73
Potential Influencer *Partisan * not Negative	1.249	0.168	7.45	<0.001	0.921	1.578
Potential Influencer * Partisan * Negative	1.899	0.295	6.43	<0.001	1.320	2.477

When a message is labelled as Bearish and contains a link to a news story from a partisan outlet, the combined effect is even bigger, as shown in Table 5(B). Specifically, the chance for it to get retweeted is 1.9%, while the chances are only 1.4% and 1.2% when the message is only labelled as Bearish or only from a partisan news source, respectively ($p < 0.001$).

As shown in Table 5(B), interestingly, the effects of message sentiment and partisan nature do not depend on the nature of the original poster of the cascades. Effects are equally sizeable whether the original poster is a non-central node or a well-connected node.

4 Social Influence of Professionals

Thus far, our investigations have shown that despite the scale-free structure, StockTwits members can mainly garner influence by attaining Official Account status and strategically utilizing message content, particularly through labeling *Negative sentiment* and referencing partisan media (see Table 5). As such, this section examines our second research question: Whether and how professionals form a subset of effective influencers.

We study the presence and effectiveness of investment advisors on StockTwits by analyzing the profiles of Official Accounts; see Section 2.3. Our observations reveal that a majority of Official Accounts are maintained by professionals, comprising 15% RIAs and 53% UnRIAs. Interestingly, professionals' followers comprise one-quarter of the entire network population, indicating their pivotal role in shaping the network structure. It might lead one to infer that professionals are effective influencers, leveraging their Official Account status to exert influence on the network and shape the flow of information. Nonetheless, two concerns arise regarding this inference, which are primarily drawn from our earlier findings in Section 3.4 concerning Official Accounts: a) We focus exclusively on the likelihood of resharing as a measure of effective influence, which may not fully capture the extent of information dissemination in the network. For instance, a cascade of size 2 (only one retweet) differs notably from a viral cascade of size 9 as shown in Fig. 2.

b) We assess the general explanatory power of Official Accounts (*Official user*) on influence without accounting for whether the account holders are professionals. To more precisely evaluate the effectiveness of professionals' use of Official Account status, we conduct a series of random-effect multinomial logistic regressions using the following equation:[14]

$$\begin{aligned} log\left(\frac{prob(Y=1)}{prob(Y=0)}\right) \\ &= \alpha + \beta_1\, Potential\ influencer + \beta_2\, Official\ user + \beta_3\, Investment\ advisor \\ &+ \beta_4\, Potential\ influencer * Investment\ advisor + \beta_5\, RIA + \beta_6\, Partisan\ nature \\ &+ \beta_7\, Sentiment + \beta_8\, Social\ attention + User + Year + \varepsilon_0.^{15} \end{aligned} \quad \textbf{(2)}$$

The equation measures the likelihood of a message going *viral* as a function of various variables, including the characteristics of the original poster, as well as the content and tone of the message. There, the outcome variable (Y), a categorical measure, reflects the *virality* of a post, quantified by the *structural virality (v)* as proposed by Goel *et al.* (2016); see Equation (1). The structural virality (v) reflects the magnitude of information dissemination in a network. Following Goel *et al.* (2016), we categorize posts into four groups based on their viral reach: Category 1 (*v < 2*), Category 2 (*broadcast: v ≈ 2*), Category 3 (*viral: v ≈ 3*), and Category 4 (*extremely viral: v > 3*). Specifically, in each regression analysis, Category 1 (small cascades) serves as Y=0 (base variable), and one of the other categories (either 2, 3, or 4) represents the measurement of Y =1. As such, each multinomial regression analysis generates three sets of results, where each of them contrasts one of the categories 2, 3, and 4 against Category 1.

α represents the random effects associated with the firms (stocks) mentioned in the messages. The regression includes poster (user) and year fixed effects. The probability of

[14] **Stata command:** xtmlogit outcome var1 var2 var3 i.user i.year || firm:, or vce(cluster user).

[15] Note that the interaction between Potential Influencer and RIA, is excluded from the regression analysis due to collinearity. This exclusion is because only 19 out of 1,585,082 observations were posted by RIAs who are not well-connected. Random effects are at the level of message author and of the firm mentioned in the message.

virality is measured conditional on structural connectivity and official status of the originator of a cascade, using two dummy variables denoting *Potential influencer* and *Official user*. In addition, we introduce two new predictor variables (compared to the resharing analysis) to assess the performance of professionals in the network, namely *Investment advisor* and *RIA*, both of which are dummy variables. The first variable indicates whether the original poster of a cascade is an investment advisor, while the second variable denotes whether the original poster is a registered investment advisor; the latter is a subset of the former. *RIA*s are a subset of investment advisors, and investment advisors themselves constitute a subset of *Official Accounts*.

We control for the tone and content, using *Sentiment* and *Partisan nature* of a message. The sentiment label assigned to a post by the original poster is regarded as the true sentiment of the post, encompassing both Bullish and Bearish perspectives. *Partisan nature* is a dummy variable equal to 1 if the original poster's message is linking to a news article published by the media deemed partisan (WSJ, NY Times, The Post, NY Post, CNN, FoxNews, and MSNBC).

We also control for *Social Attention*, which is defined as the aggregate number of *Likes* received by the entire cascade, scaled by the original poster's number of followers: *Social Attention = Likes / (Original tweeter's followers + 0.5).* This control aims to capture the unobserved effects of message content not explained by *Sentiment* and *Partisan nature*. For instance, a message may attain high Social Attention and virality by imparting novel and informative information about a stock.

Results are presented in Table 6 across three models. In Model 1, Equation (2) is employed, excluding *Social Attention* and the interaction between *Potential influencer* and *Investment advisor*. Model 2 uses Equation (2) with *Social Attention* excluded. Model 3 provides a more comprehensive specification by including all explanatory variables outlined in Equation (2). Here, we focus on the regression analysis results using Model 3 as it has the lowest *Akaike Information Criterion* (AIC) and highest R^2.

Table 6: Effectiveness of Professional Investment Advisors

This table shows the results of a series of random-effects multinomial logistic regression analyses (panel variables: firms mentioned in messages) across three models, using Equation (2). The equation measures the likelihood of a message going *viral* as a function of various variables, including the characteristics of the original poster, as well as the content and tone of the message. The outcome variable is *structural virality (v)*, assessed across four categories: Category 1 ($v < 2$), Category 2 (broadcast: $v \cong 2$), Category 3 (viral: $v \cong 3$), and Category 4 (extremely viral: $v > 3$). Specifically, in each regression analysis, Category 1 (small cascades) serves as Y=0 (base variable), and one of the other categories (either 2, 3, or 4) represents the measurement of Y =1. Original poster characteristics are measured through *Potential influencer*, *Official user*, *Investment advisor*, and *RIA*. *Potential influencer* is a dummy variable equal to 1 if the original poster does have more than 1000 followers. *Official user* is a dummy variable equal to 1 if the original poster holds an *Official Account*. *Investment advisor* is a dummy variable equal to 1 if the original poster holds an Official Account and is a professional in the investment field associated with investment advice companies or websites providing investment tools. *RIA* is a dummy variable set to 1 if the original poster holds an *Official Account* and is a registered investment advisor. *Social Attention* is equal to the number of *Likes* received by the cascades scaled by the original poster's number of followers: *Social Attention = Likes / (Original tweeter's followers + 0.5)*. *Sentiment* is a categorical variable equal to 1 and 2 if the original poster has labeled his/her post as *Bullish* and *Bearish*, respectively. *Sentiment* equal to 0 if the original poster has not labeled his/her post. *Partisan nature* is a dummy variable equal to 1 if the original poster's tweet linking to a news article published by the media deemed partisan (WSJ, NY Times, The Post, NY Post, CNN, FoxNews, and MSNBC). All models include poster and year fixed effects. The odds ratios (exponentiated coefficients) for β_1 to β_8 are reported across all regressions. Standard errors, clustered at the user (poster) level, are reported in parentheses. *, **, and *** denote significance at the 0.05, 0.01, and 0.001 levels, respectively.

Model	**(1)**			**(2)**			**(3)**		
Depth Category (i)	2	3	4	2	3	4	2	3	4
Potential influencer	0.683***	0.722***	0.804**	0.691***	0.735***	0.823**	0.721***	0.775***	0.867*
	(0.008)	(0.028)	(0.059)	(0.009)	(0.029)	(0.061)	(0.009)	(0.029)	(0.062)
Official user	1.406***	1.496***	1.686***	1.393***	1.473***	1.657***	1.113***	1.345***	1.632***
	(0.029)	(0.093)	(0.187)	(0.029)	(0.092)	(0.184)	(0.029)	(0.091)	(0.183)
Investment advisor	1.041	1.043	2.868***	1.044	1.047	2.885***	0.288***	0.000	0.599
	(0.023)	(0.069)	(0.299)	(0.023)	(0.069)	(0.300)	(0.065)	(0.000)	(0.599)
Potential influencer # Investment advisor							4.806***	5.687**	7.364*
							(1.084)	(2.331)	(7.390)
RIA	3.571***	4.707***	1.421*	3.575***	4.713***	1.425*	3.542***	4.662***	1.413*
	(0.119)	(0.415)	(0.225)	(0.119)	(0.416)	(0.226)	(0.118)	(0.411)	(0.224)
Partisan nature	1.619***	1.659***	1.439	1.607***	1.644***	1.415	1.615***	1.654***	1.421
	(0.061)	(0.184)	(0.345)	(0.060)	(0.183)	(0.340)	(0.061)	(0.184)	(0.341)
Negative sentiment	1.913***	2.430***	2.099***	1.887***	2.387***	2.056***	1.872***	2.362***	2.024***
	(0.024)	(0.091)	(0.142)	(0.024)	(0.089)	(0.140)	(0.024)	(0.088)	(0.137)
Positive sentiment	2.145***	2.735***	1.909***	2.157***	2.766***	1.934***	2.142***	2.739***	1.909***
	(0.047)	(0.171)	(0.235)	(0.047)	(0.173)	(0.238)	(0.047)	(0.172)	(0.235)
Social attention				1.148***	1.179***	1.209***	1.149***	1.181***	1.211***
				(0.005)	(0.009)	(0.012)	(0.005)	(0.009)	(0.012)
Constant	0.0031***	0.0004***	0.0002***	0.0031***	0.0004***	0.0002***	0.0032***	0.0004***	0.0002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations		1,585,082			1,585,082			1,585,082	
AIC		489638.3			488726.6			487321.2	
χ^2		8979.779			9897.484			11308.878	

Notably, as shown by the odds ratios for *Constant*, broadcast ($v \cong 2$), viral ($v \cong 3$), and extremely viral ($v > 3$) cascades constitute about 32 bps, 4 bps, and 2 bps of the total observations (1,585,082 cascades), respectively.

Results suggest that potential influencers attain a lower level of virality within the network when contrasted with non-central members. The probability of a message from a potential influencer achieving virality in the network is approximately 28%, 22%, and 13% *lower* than that from a non-central member for virality of 2 (broadcast), 3 (viral), and over 3 (extremely viral), respectively. The results are highly significant across the three models (p<0.001). This strongly supports our earlier proposition, indicating that *structural connectivity* does not contribute to information dissemination in the network, thus, to *functional connectivity*; see Sections 3.3 and 3.4.

In contrast, well-connected investment advisors measured by the interaction term *Potential influencer * Investment advisor*, generate significantly more viral messages than regular users (baseline). The likelihood of their posts going viral is 381%, 469%, and 636% *higher* than that for regular users at virality of 2 (broadcast), 3 (viral), and over 3 (extremely viral), respectively.

Notably, the effect of RIAs is even higher than that for general investment advisors (note that RIA is a subset of Investment advisor). The likelihood of viral messages from RIAs is 254%, 366%, and 41% higher than that for general investment advisors at the structural virality of 2, 3, and over 3, respectively. This is not only sizeable, but also highly significant (p< 0.001). Given the majority of RIAs being well-connected, the lack of influence among potential influencers suggests that the effect of RIAs extends beyond the effects of structural connectivity and general investment advisors.

Consistent with earlier findings (see Section 3.4), both *Partisan nature* and Bullish/Bearish *Sentiment* amplify the probability of a message going viral. Specifically, referencing partisan news outlets increases the likelihood of virality by 62% and 65% for broadcast (v≅ 2) and viral

($v \cong 3$) posts, respectively. Bearish (Negative) self-labeled sentiment enhances that likelihood by 87%, 136%, and 102% at virality of 2 (broadcast), 3 (viral), and over 3 (extremely viral), respectively. Bullish sentiment elevates the likelihood of virality by 114%, 174%, and 91% at virality of 2, 3, and over 3, respectively.

The results remain consistent across models 1 and 2. Specifically, as evidenced by Category 4 in Model 1, RIAs stand out as the most influential among various Official Account groups on StockTwits. There, Official Accounts have a 69% higher chance of their messages going viral compared to regular users. This likelihood increases by 187% when the Official Account holder is also an investment advisor. The likelihood even rises by an additional 42% when an investment advisor is also registered with either the SEC or state securities administrators.[16]

We conduct additional multinomial logistic regression analysis to assess the *Social Attention* garnered by professional investment advisors, measured by the number of *likes* received from other users. The results of the regression analysis are presented in Table A. 2 in the Appendix. Our findings indicate that professionals, particularly RIAs, distinguish themselves by garnering notably higher influence and Social Attention in comparison to other users on StockTwits.

As a robustness test, we estimate an OLS model with three-way fixed effects for time, user (poster), and firm (stocks mentioned in the post).[17] The results, reported in **Table 7**, are consistent with our multinomial logistic regression findings. The coefficient on RIAs ($\beta_2 = 0.071$, $t = 14.30$, $p < 0.001$) indicates that, controlling for user, firm, and year fixed effects, posts by registered investment advisors attain significantly higher virality scores than those by non-official users. This evidence suggests that professional status enhances information diffusion and audience engagement on the platform.

[16] Note that RIAs are a subset of investment advisors, and investment advisors themselves constitute a subset of official accounts.

[17] **Stata command:** reghdfe outcome var1 var2 var3, absorb (user firm year) vce (cluster User).

Table 7: OLS Model with Three-Way Fixed Effects

This table reports the results of the OLS model examining the influence of registered (RIAs) and unregistered (UnRIAs) investment advisors on StockTwits, estimated using a three-way fixed-effects specification (number of observations = 1,585,082 and Adjusted R^2 = 0.1217):

$$y_{imt} = \alpha + \beta_1 \text{Potential Influencer}_{imt} + \beta_2 \text{Official User}_{imt} + \beta_3 \text{Partisan Nature}_{imt} + \beta_4 \text{Sentiment}_{imt} + \text{UserFE}_i + \text{FirmFE}_m + \text{YearFE}_t + \varepsilon_{imt}.$$

The dependent variable y_{imt} is the virality category of a message posted by user i on stock m in time t, taking values from 1 (non-viral) to 4 (extremely viral). Official user equals 1 for unregistered advisors, 2 for registered advisors, and 0 otherwise. Potential influencer is a dummy variable equal to 1 if the original poster does have more than 1000 followers. Sentiment is a categorical variable equal to 1 and 2 if the original poster has labeled his/her post as Bullish and Bearish, respectively, 0 otherwise. Partisan nature is a dummy variable equal to 1 if the original poster's tweet linking to a news article published by the media deemed partisan (WSJ, NY Times, The Post, NY Post, CNN, FoxNews, and MSNBC). The model includes user, firm (stock mentioned in the post), and year fixed effects to control for unobserved heterogeneity. Standard errors, clustered at the user level.

	Coefficient	Robust std. err.	t	P> |t|	[95% conf.	interval]
Potential influencer	0.014	0.023	0.63	0.526	-0.030	0.059
Official Accounts						
UnRIA	-0.023	0.090	-0.25	0.800	-0.199	0.154
RIA	0.071	0.005	14.3	0.000	0.061	0.080
Partisan nature	-0.001	0.003	-0.33	0.744	-0.007	0.005
Sentiment						
Bullish	0.008	0.002	4.35	0.000	0.004	0.012
Bearish	0.025	0.003	8.38	0.000	0.019	0.031
Constant	1.024	0.019	53.43	0.000	0.987	1.062

4.1 Who Exploits Partisan Nature and Sentiment?

We found earlier that message content, specifically partisan nature and sentiment, appear to matter as much as location and status for influence. Thus, it would be intriguing to investigate individuals who exploit partisan references and sentiment to gain influence within the network. To assess the role of message content on the influence gained by StockTwits members, we run two logistic regressions with firm specific random effects. The first logistic regression a random-effects multinomial logistic regression analysis, in which the outcome variable (Y) is categorical with three values: Category 1 (no self-declared sentiment label, serving as the base), Category 2 (Bullish), and Category 3 (Bearish). In each regression, Category 1 is treated as Y=0, and either Category 2 or 3 is compared as Y=1. In the second logistic regression, the dependent variable is whether one's message refers to a story from a partisan news outlet (*Partisan nature*); yes = 1; no = 0.

In both models, the key explanatory variables capture characteristics of the original poster: "Potential Influencer" (a dummy equal to 1 if the user has >1000 followers) and "User Type" (RIAs, UnRIAs, other Official Accounts (Other OAs), and regular users as the base). Both models include poster and year fixed effects. The results of the multinomial and binary logistic regressions are presented in Table 8 and Table 9, respectively.

The most striking result from this study is that the least connected nodes exploit the enhanced impact of negative sentiment and partisan news for gaining influence within the network. Specifically, regarding Bearish (Negative) labeling, the likelihood of negative labeled messages from a potential influencer is ~88% less than that for a less-connected member (significant at $p<0.001$); see Table 8. Concerning *Partisan nature*, the likelihood of one's message referring to a story from a partisan news outlet also decreases by ~82% when the post is authored by a potential influencer compared to a poorly connected member (significant at $p<0.001$); see Table 9.

Fig. A. 3 in the Appendix also represents how least connected members strategically combine negative sentiment and partisan nature in their communications to maximize their advantage from these two factors.

When posting about a stock like Apple Inc., we observed that referencing a partisan news story—often with a reframed narrative—is an effective way to attract attention. For example, a news article favorable to Apple but unfavorable to Google may be presented as a Bearish message about Google, while still mentioning Apple.

Table 8: Impact of Originator Identity on Message Sentiment

This table presents the outcomes of a random-effects multinomial logistic regression analysis (panel variables: firm), estimating the likelihood of labeling a message as Positive (Bullish) or Negative (Bearish), as a function of variables including the characteristics of the author of the message using the following equation (Number of observations = 1,585,082, χ^2 = 166,848, Log likelihood = -739,275):

$$\log\left(\frac{prob\ (Y=1)}{prob\ (Y=0)}\right) = \alpha + \beta_1\ Potential\ influencer + \beta_2\ User\ type + User + Year + \varepsilon_0\ .$$

The outcome variable (Y) is categorical with three values: Category 1 (no self-declared sentiment label, serving as the base), Category 2 (Bullish), and Category 3 (Bearish). In each regression, Category 1 is treated as Y=0, and either Category 2 or 3 is compared as Y=1. The model includes random effects α at the firm level (stocks mentioned in the message) and predictors related to the original tweeter, including "Potential Influencer" (a dummy equal to 1 if the user has >1000 followers) and "User Type" (RIAs, UnRIAs, other Official Accounts (Other OAs), and regular users as the base). The regression includes poster and year fixed effects. The odds ratios for β_1 and β_2 are reported. The ***constant*** captures baseline relative risk, conditional on zero random effects. A likelihood ratio test comparing this model to a standard multinomial logit yields **χ2(2)=2,675** (p<0.0001), justifying the panel structure. The estimated variance components show that **var(u1)=1.995** (for Bullish) and **var(u2)=0.63** (for Bearish), indicating unobserved firm-level heterogeneity in sentiment outcomes.

Y=1	Predictor	Odds Ratio	Std. Err.	Z	p-value	95% Confid. Interval	
	Potential Influencer	0.122	0.001	-333.31	<0.001	0.120	0.123
	User Type						
Sentiment	Other OA	0.346	0.007	-49.51	<0.001	0.332	0.361
Category = Bullish	UnRIA	1.255	0.013	21.71	<0.001	1.229	1.281
	RIA	0.180	0.016	-19.08	<0.001	0.151	0.215
	Constant	0.469	0.001	-264.92	<0.001	0.466	0.472
	Potential Influencer	0.150	0.002	-147.66	<0.001	0.146	0.154
Sentiment	**User Type**						
Category = Bearish	Other OA	0.595	0.021	-14.70	<0.001	0.555	0.638
	UnRIA	2.203	0.039	44.38	<0.001	2.127	2.281
	RIA	0.394	0.051	-7.22	<0.001	0.306	0.508
	Constant	0.084	0.000	-428.50	<0.001	0.083	0.085
Var(u1)		1.995	0.412			1.331	2.991
Var (u2)		0.630	0.148			0.397	0.999

More importantly, we also find that RIAs' influence is not derived from the partisan nature or Bullish/Bearish tone labeling of their messages. However, it is observed that UnRIAs have effectively used *sentiment* to establish influence within the network.

As shown in Table 8, RIAs are significantly less likely to label their posts as Bullish or Bearish—by approximately 82% and 61%, respectively—compared to regular users. In contrast, UnRIAs are substantially more likely to use sentiment labels, with the odds of Bullish and Bearish labeling increasing by approximately 255% and 203%, respectively.

Table 9 further shows that both RIAs and UnRIAs are less likely than regular users to reference stories from partisan news outlets, by approximately 34% and 50%, respectively.

Table 9: Impact of Originator Identity on Partisan Nature

This table reports the results of a random-effects logistic regression (panel variable: firm), estimating the likelihood of referencing a story from a partisan news outlet ("Partisan nature") as a function of message sentiment and author characteristics, based on the following equation (Number of observations = 1,585,082, χ^2 = 8,604, Log likelihood = -69,837):

$$\log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 Potential\ influencer + \beta_2\ \text{User type} + \beta_3\ Sentiment + \text{User} + \text{Year} + \varepsilon_0\ ,$$

p denotes the probability of posting a message referring to partisan media (WSJ, NY Times, The Post, NY Post, CNN, Fox News, MSNBC). The model includes random intercepts α at the firm level (i.e., stocks mentioned in the message) and predictors capturing characteristics of the original poster and the tone of the message: **Potential Influencer** (a dummy equal to 1 if the user has more than 1,000 followers), **User Type** (RIAs, UnRIAs, other official accounts [Other OAs], with regular users as the reference category), and **Sentiment**, a categorical variable reflecting the self-declared tone of the message (Without label, Bullish, Bearish). The model controls for user and year fixed effects. The odds ratios for β_1 to β_3 are reported. A likelihood ratio test comparing this model to a standard logit yields $\chi2(2)$ = 4,814 (p<0.0001), justifying the panel structure.

Predictor	Odds Ratio	Std. Err.	Z	p-value	95% Confid. Interval	
Potential Influencer	0.179	0.005	-62.23	<0.001	0.170	0.190
User Type						
Official user	1.179	0.069	2.79	0.005	0.332	0.361
UnRIA	0.502	0.033	-10.39	<0.001	1.050	1.323
RIA	0.661	0.057	-12.95	<0.001	0.398	0.890
Sentiment						
Bullish	1.380	0.028	15.66	<0.001	1.326	1.437
Bearish	3.322	0.093	43.11	<0.001	3.146	3.509
Constant	0.004	0.001	-38.66	<0.001	0.003	0.005
var(Constant)	**1.763**	**0.344**			**1.202**	**2.585**

4.2 Temporal Variations of Professionals' Network Activities

Our findings in the previous section highlight the considerable influence of professionals, particularly RIAs, within the network, displaying statistically significant and economically sizable results. This might lead to the expectation of investment advisors, specifically RIAs, progressively leveraging social media for their business purposes. Surprisingly, as discussed in this section, our findings indicate the opposite!

To explore this proposition, we study the temporal variations of professionals' network activities over our study time span, from January 2014 to December 2020. This investigation encompasses the evaluation of data traffic, specifically the frequency of posts, and the count of active investment advisors within the network. Surprisingly, there was a noteworthy decline in professionals' posts starting from 2018; see Fig. 9. While the aggregated monthly message counts from other members surged thirteenfold to 325,000 in December 2020 from 25,000 in January 2015, the count for RIAs decreased from 400 in January 2015 to one-quarter of that number (100) in December 2020.

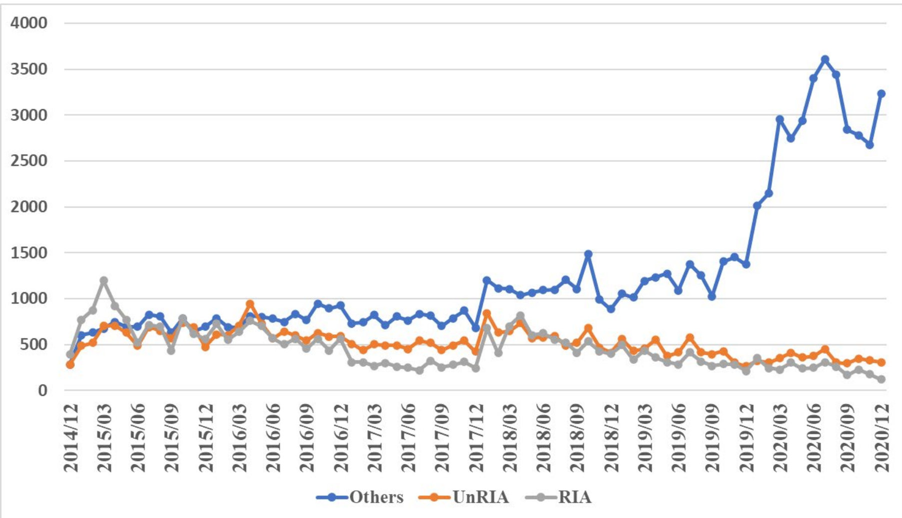

Fig. 9: Post Dynamics on StockTwits

Figure depicts the aggregated monthly message counts posted by StockTwits users from January 2014 to December 2020, categorized into three groups: RIAs, UnRIAs, and other users. The message counts of UnRIAs and other users are divided by 10 and 100, respectively, for visualization purposes.

Additionally, on the one hand, as indicated in Panel A of Table 10, there was a substantial decline in the addition of new investment advisors (both RIAs and UnRIAs) to the network, from 2018 to 2020. On the other hand, as revealed in Panel B of Table 10, a considerable number of investment advisors discontinued their usage of StockTwits during that period. These dual factors contributed to a progressive decrease in the number of active investment advisors in the network.

Table 10: Official Account Dynamics on StockTwits: New Additions and Abandonments

This table presents the number of Official Account additions (Panel A) and abandonments (Panel B) on StockTwits from January 2009 to December 2020.

Panel A) Frequency of StockTwits' new official accounts

User class	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total
RIAs	9	11	14	8	13	4	5	1	3	2	1	1	72
UnRIAs	27	46	41	27	25	36	20	20	5	8	3	2	260
Others	9	20	26	6	20	25	10	5	6	2	0	0	129
Total	45	77	81	41	58	65	35	26	14	12	4	3	461

Panel B) Frequency of StockTwits' abandoned official accounts

User class	2015	2016	2017	2018	2019	2020	Total
RIAs	6	5	7	5	5	9	37
UnRIAs	7	13	17	17	30	25	109
Others	9	13	13	22	9	14	80
Total	22	31	37	44	44	48	226

Consequently, we extend our research to ascertain additional evidence supporting our observation regarding the reluctance of investment advisors to use social media since 2018. Our findings include a few news articles and a survey that align with this trend. In particular, a warning article published by K&L Gates on July 17, 2018, cautions investment advisors against social media postings, highlighting the SEC's ongoing scrutiny.[18] The Putnam Social Advisor also conducted a survey in 2023, revealing that numerous investment advisors have

[18] Available online at Be Careful What you Post - SEC Continues To Focus on the Use of Social Media by Investment Advisers | HUB | K&L Gates (klgates.com).

refrained from using social platforms like Facebook and Twitter for business due to compliance policies. The survey involved 1,043 U.S. financial advisors with over two years of experience advising retail clients.[19]

5 SEC Regulations and Investment Advisors' use of social media

Our findings indicate that social media has become integral to investment advisors' practices, facilitating retail client engagement and brand expansion. The SEC and FINRA govern the conduct of investment advisors, particularly RIAs. Since 2017, these regulatory bodies have intensified their oversight of advisors' use of social media to protect retail investors. Recent enforcement actions raise concerns about potential unintended consequences, as noted in K&L Gates' investment management alert on advisors' use of social media. Overregulation or ineffective communication by the SEC may lead to a *chilling effect*, wherein the fear of SEC enforcement leads advisors to excessive self-censorship, ultimately paralyzing professional insights on social media.

This section examines whether regulatory oversight yields unintended consequences, such as a chilling effect. We analyze advisors' postings and followings on social media from January 2014 to December 2020. Given that SEC enforcement began in July 2018, we compare investment advisors' influence and engagement on social media before and after its implementation. During the enforcement period of our study (July 2018–December 2020), the SEC and FINRA actively monitored advisors' social media practices to ensure fair disclosure, prevent violations, and detect market manipulation (e.g., pump-and-dump schemes) to protect retail investors. This oversight has led to enforcement actions related to advisors' use of social media, including several regulatory notices, amendments to SEC Form ADV and related rules, and SEC fines. For example, a recent SEC Press-Release announced $400,000 in fines for

[19] Available online at https://www.putnam.com/static/pdf/Putnam-Social-Advisor-Survey-2023.pdf.

misleading claims on social media as part of settled charges against two registered investment advisors. While the SEC and FINRA have stated that they "seek to interpret their rules in a flexible manner to allow advisors to communicate with clients and investors using social media (see FINRA, 2010)," the prospect of regulatory scrutiny may nonetheless had a chilling effect on advisors' online engagement.

5.1 Impact of SEC Restrictions: Background

The primary FINRA regulation concerning blogs and social networking websites was issued on January 25, 2010, through Regulatory Notice 10-06.[20] It states that social networking sites like Facebook, Twitter, and LinkedIn typically feature both static and interactive content. Static content (such as profile, background, or wall information) falls under Rule 2210, requiring approval from a registered principal before posting on a firm's or representative's social networking page. Conversely, non-static, real-time communications like interactive posts on platforms like Twitter and Facebook do not require prior approval from a registered principal. However, firms are still responsible for supervising these communications with respect to FINRA's communication rules, namely Rule 2111 and Rules 2210 through 2220. These rules necessitate that advisors ensure their social media communications are fair, balanced, and free from misleading information.[21] Moreover, in March 2014, the SEC published guidance that encourages investment advisors to be cautious of favorable social media commentary that may potentially violate SEC Rule 206(4)-1(a)(1) in their advertisements.[22]

While these rules govern the use of social media by investment advisors, they do not outright restrict RIAs' activity on social media platforms. However, starting from October 2016 until November 2019, there was an increasing focus from both the SEC and FINRA on monitoring

---

[20] Available online at https://www.finra.org/rules-guidance/notices/10-06

[21] FINRA Rule 2111 mandates that member firms and associated individuals must have a reasonable foundation to believe that a recommended transaction or investment strategy involving a security is suitable for customers.

[22] Available online at https://www.sec.gov/investment/im-guidance-2014-04.pdf.

the use of social media by investment advisors. As a result of this heightened scrutiny, certain limitations were introduced, and the SEC began actively enforcing regulations more rigorously.

On October 26, 2016, FINRA published Regulatory Notice 16-41, announcing that the SEC had approved amendments to rules governing communications with the public, which would be effective from January 9, 2017.[23] The notice explains that FINRA adopted amendments to the filing requirements in FINRA Rules 2210 and 2214, specifically targeting providers of investment analysis tools. This regulatory notice was significant for many StockTwits RIAs who offer investment analysis tools and utilize social media for advertising purposes. This also posed a threat to StockTwits UnRIAs.

On April 25, 2017, FINRA published Regulatory Notice 17-18, providing guidance on social networking websites and business communications.[24] This guidance also imposes unfavorable restrictions on investment advisors (i.e., both RIAs and UnRIAs). The notice states that a firm has adopted content by sharing or linking to specific content posted by an independent third party. Therefore, the firm would be responsible for ensuring that the content complies with the same standards as communications created by, or on behalf of, the firm. This is, thus, subject to the communications rules (i.e., FINRA Rule 2111), including the prohibition on misleading or incomplete statements or claims, the testimonial requirements (i.e., SEC Rule 206(4)-1), and the supervision and recordkeeping rules (i.e., Rules 2210 through 2220).

Starting October 1, 2017, the SEC began requiring all RIAs to provide additional information on their Form ADV disclosing all separately managed social media accounts.[25] This includes details of all their social media accounts such as Facebook, Twitter/StockTwits, and LinkedIn, as well as the advisor's website URL.[26]

---

[23] Available online at https://www.finra.org/rules-guidance/notices/16-41.
[24] Available online at https://www.finra.org/rules-guidance/notices/17-18.
[25] Form ADV is the uniform form used by investment advisors to register with both the Securities and Exchange Commission (SEC) and state securities authorities.
[26] Available online at https://www.sec.gov/news/press-release/2016-168.

On July 10, 2018, the SEC reaffirmed the application of securities laws to social media use through five settlements. The settlements involved RIAs and their representatives, as well as a marketing consultant, who violated the SEC Testimonial Rule, Rule 206(4)-1(a)(1), by publishing client testimonials on social media and other websites. These settlements demonstrate the SEC's active enforcement of the Testimonial Rule and increased focus on social media use by RIAs.[27]

On November 4, 2019, the SEC published amendments to Rule 206(4)-1 (the "Advertising Rule") and Rule 206(4)-3 (the "Solicitation Rule") under the Investment Advisors Act of 1940. There, the definition of "advertisement" was broadened to include any type of communication including social media.[28]

5.2 Impact of SEC Restrictions: Natural Experiment

In light of the SEC's regulatory context, we conduct a series of Difference-in-Differences (DiD) analyses—encompassing both standard and dynamic two-way fixed effects (TWFE) specifications with time and user fixed effects—to assess the impact of SEC restrictions on user engagement on StockTwits. Initially, we evaluate the impact of the SEC enforcement on the posting activity of investment advisors and other users using the following specification for the standard DiD model:

$$y_{it} = \alpha + \beta D_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \qquad i = 1, \dots, N; \quad t = 1, \dots, T. \tag{3}$$

Users $i = 1, \dots, N = 69{,}314$ (control $n = 68{,}866$; treated $n = 448$); months $t = 1, \dots, T$ with $T = 84$ spanning January 2014–December 2020. The SEC policy event is dated **July 2018** ($t = 54$), when the Commission reaffirmed the application of federal securities laws to professionals' use of social media and announced five settlements. The outcome y_{it} is the

[27] Available online at https://www.klgates.com/Be-Careful-What-you-Post---SEC-Continues-To-Focus-on-the-Use-of-Social-Media-by-Investment-Advisers-07-17-2018.

[28] Available online at SEC_Proposes_Amendments_to_Update_the_Investment_Adviser_Advertising_and_Solicitation_Rules.pdf (willkie.com).

logarithm of the number of messages (logMessage) posted by user i in period t, aggregated monthly from January 2014 to December 2020. The model includes time fixed effects λ_t, user fixed effects μ_i, and an error term ε_{it}. The treatment variable D_{it} is defined as either (i) a categorical variable equal to 1 for UnRIAs, 2 for RIAs in periods t after the SEC event, and 0 otherwise; or (ii) a dummy variable equal to 1 for professionals (advisors) in periods t after the SEC event, and 0 otherwise. We restrict the sample to users with established posting histories on StockTwits before the SEC event and construct a balanced panel by coding missing values of y_{it} as zero, since the absence of a record on a given day indicates no posts related to the relevant stocks.[29]

Models 1–2 and 4-5 in Table 11 present the results for the TWFE DiD analysis based on Equation (3). As shown in Model 2, the average treatment effect on the treated (ATT) for UnRIAs is −0.549, corresponding to a 42.2% decline (exp(−0.549) − 1) in message counts for UnRIAs relative to other StockTwits users (the control group), statistically significant at the 0.001 level. This decline further intensifies to 53.0% (exp(−0.756) − 1) for RIAs, also significant at the 0.001 level. Notably, prior to the SEC treatment, UnRIAs and RIAs posted 678% (exp(2.052) − 1) and 487% (exp(1.769) − 1) more messages, respectively, than other members (see Model 1). Moreover, the overall ATT for professionals is −0.613 (see Model 4). These findings demonstrate that the SEC restrictions had a substantial chilling effect on the use of StockTwits by investment advisors, particularly RIAs. Following the implementation of these restrictions, investment advisors experienced a significant and pronounced decline in their monthly message activity compared to other network members, despite their previously high levels of engagement in both posting activity and network influence.

Model 3 and 6 in Table 11 report the post-SEC ATT estimates obtained from the inverse propensity score weighted (IPW) DiD approach of Callaway and Sant'Anna (2021), estimated

[29] **Stata command:** reghdfe logMessage i.treated##after, absorb (UserID mdate) vce(cluster UserID).

using Equation (3). Model 6 additionally presents the pre-SEC ATT to assess the presence of treatment effects prior to the intervention (i.e., the *no-anticipation assumption*). Integrating IPW into the TWFE specification supports the conditional parallel trends assumption for identification.

Table 11: Impact of SEC Restrictions (Difference-in-Differences)

This table reports difference-in-differences estimates from six models using individual-level panel data, using Equation (3). These models incorporate time (months $t = 1, \ldots, T$, with $T = 84$ spanning January 2014–December 2020) and user fixed effects (control group: $n = 68{,}866$; treated group: $n = 448$). The outcome variable is the logarithmically transformed number of monthly messages posted by a StockTwits user, defined as "log(monthly message volume + 0.5)." The SEC policy event is dated July 2018 (t = 54), when the Commission reaffirmed the applicability of federal securities laws to professionals' use of social media and announced five settlements. The treatment variable is defined as either (i) a categorical variable equal to 1 for UnRIAs, 2 for RIAs in periods t after post-SEC, and 0 otherwise (Models 1-3); or (ii) a dummy variable equal to 1 for professionals (advisors) in periods t after post-SEC, and 0 otherwise (Models 4-6). Models 1–2 and 4-5 report the ATT estimates from the standard TWFE specification. Models 3 and 6 present ATT estimates from the inverse propensity score weighting (IPW) DiD of Callaway and Sant'Anna (2021). The standard DiD models compare treated users to all controls with equal weights, without ensuring covariate balance, whereas the IPW approach (Models 3 and 6) reweights controls to enhance comparability. Refer to the IPW DiD section in Appendix for a detailed explanation of the IPW DiD model. In all models, standard errors are clustered at the user level (69,314 clusters) with t-statistics shown in parentheses. *, **, and *** denote significance at the 0.05, 0.01, and 0.001 levels, respectively.

Model	(1) DiD	(2) DiD	(3) IPW DiD	(4) DiD	(5) DiD	(6) IPW DiD
RIA	1.769***					
	(9.52)					
UnRIA	2.052***					
	(17.24)					
ATT (RIA)	- 0.756***	-0.756***	-0.337*			
	(-5.27)	(-5.27)	(-2.23)			
ATT (UnRIA)	-0.549***	-0.549***	-0.131			
	(-5.55)	(-5.55)	(-1.19)			
Advisors				1.964***		
				(19.54)		
ATT (Advisors_Post_SEC)				-0.613***	-0.613***	-0.195*
				(-7.51)	(-7.51)	(-2.05)
ATT (Advisors_Pre_SEC)						-0.008
						(-1.69)
Time fixed effects	No	Yes	Yes	No	Yes	Yes
User fixed effects	No	Yes	Yes	No	Yes	Yes
Observations	5,059,922	5,059,922	5,059,922	5,059,922	5,059,922	5,059,922
Adjusted R-squared	0.0143	0.2710	0.4576	0.0142	0.2710	0.4576

The IPW procedure reweights observations by the inverse of their treatment propensity, generating a pseudo-population that approximates randomized treatment assignment and achieves covariate balance between treated and control units, thereby supporting the conditional parallel trends assumption. This covariate-balanced pseudo-population further facilitates the estimation of group–time average treatment effects while mitigating confounding. In our study, the use of IPW is particularly relevant given the pronounced sample imbalance in our data (control $n = 68{,}866$; treated $n = 448$). In this context, average propensity scores during the pre-treatment period are estimated based on covariates capturing users' network characteristics prior to the SEC event. These covariates include the number of Followers, Following, Watchlist Size, Ideas, Membership Days, and Likes per user. [30] Refer to the IPW DiD section in Appendix for a detailed description of the IPW DiD methodology.[31]

According to Model 3 and 6 in Table 11, the ATT for RIAs estimated using the IPW DiD approach (Model 3, −0.337) is statistically significant at the 5% level and approximately half the magnitude of the standard TWFE estimate (Model 2, −0.756). In contrast, the IPW DiD estimate of the ATT for UnRIAs is statistically insignificant at the 5% level. Furthermore, the overall ATT for professionals (both registered and unregistered) obtained from the IPW DiD model (Model 6, −0.195) is statistically significant at the 5% level and roughly one-third the magnitude of the corresponding TWFE estimate (Model 4, −0.613).

These results suggest that the standard TWFE specification likely overstates the impact of the SEC policy changes on professionals. This discrepancy arises because the conventional DiD assigns equal weights to all control observations without ensuring covariate balance,

---

[30] **Followers** represent the average number of followers of user *i* during month *m*; **Following** denotes the average number of accounts followed by user *i* during month *m*; **Watchlist Size** is the average number of securities on user *i*'s watchlist during month *m*; **Ideas (cumulative posts)** indicate the total number of messages posted by user *i* from their join date through the end of month *m*; and **UserLikes (cumulative likes)** capture the total number of likes received by user *i* from their join date through the end of month *m* in year *y*. **Membership days** denotes the number of days since user i's join date through the end of month m in year y.

[31] **Stata command:** csdid covariates, ivar(userID) time(mdate) gvar(g) method(ipw); or, xtdidregress (logMessage) (did) [pweight= pscore], group(UserID) time(mdate) vce(cluster UserID).

whereas the IPW DiD enhances comparability between treated and control groups by conditioning on observed covariates and reweighting observations. Consequently, the IPW DiD provides stronger support for the validity of the parallel trends assumption. Moreover, the pre-policy ATT from the IPW DiD (Model 6) is close to zero and statistically insignificant at the 5% level, supporting the model's validity under the standard *no-anticipation assumption*. Overall, the IPW DiD estimates corroborate the chilling effect of the SEC restrictions on RIAs, albeit through a more conservative and methodologically robust approach compared to the standard TWFE specification.

Next, we examine the dynamic effects of the SEC policy on professionals' posting activity using IPW and the following specification for the dynamic DiD analysis:

$$y_{it} = \alpha + \sum_{j=2}^{J} \beta_j \, (Lead\, j)_{\,it} + \sum\nolimits_{k=1}^{K} \gamma_k \, (lag\ k)_{it} + \mu_i + \lambda_t + \varepsilon_{it}. \quad (4)$$

The dynamic DiD specification includes leads and lags of the SEC event: the leads assess the validity of the no-anticipation assumption by testing for pre-treatment dynamics, while the post-treatment lags capture the temporal pattern of the policy's impact, indicating whether the effects strengthen, attenuate, or persist over time. Specifically, leads (pre-treatment indicators) and lags (post-treatment indicators) are dummy variables. For never-treated (control) users, all these dummies are always 0. For treated users (professionals) these dummy variables are defined as follows:

$$(\text{Lead J})_{\text{it}} = 1[\text{t} \le \text{Event} - \text{J}], \quad (5)$$

$$(\text{Lead j})_{\text{it}} = 1[\text{t} = \text{Event} - \text{j}] \quad \text{for j} = 1, \ldots, \text{and J} - 1, \quad (6)$$

$$(\text{Lag k})_{\text{it}} = 1[\text{t} = \text{Event} + \text{k}] \quad \text{for k} = 1, \ldots, \text{and K} - 1, \quad (7)$$

$$(\text{Lag K})_{\text{it}} = 1[\text{t} \ge \text{Event} + \text{K}]. \quad (8)$$

Where Event denotes the date of the SEC policy change—the introduction of enforcement actions supporting the amended regulations in July 2018. In our setting, all treated users (professionals) share the same event date, as the policy was implemented uniformly rather than in a staggered fashion. As such, non-treated users act as the reference group via time fixed effects, while the lead/lag dummies capture deviations for professionals around the policy event. As standard, the baseline omitted period is the first lead, where J= 1. Fig. 10 presents the results for the dynamic IPW-DiD.

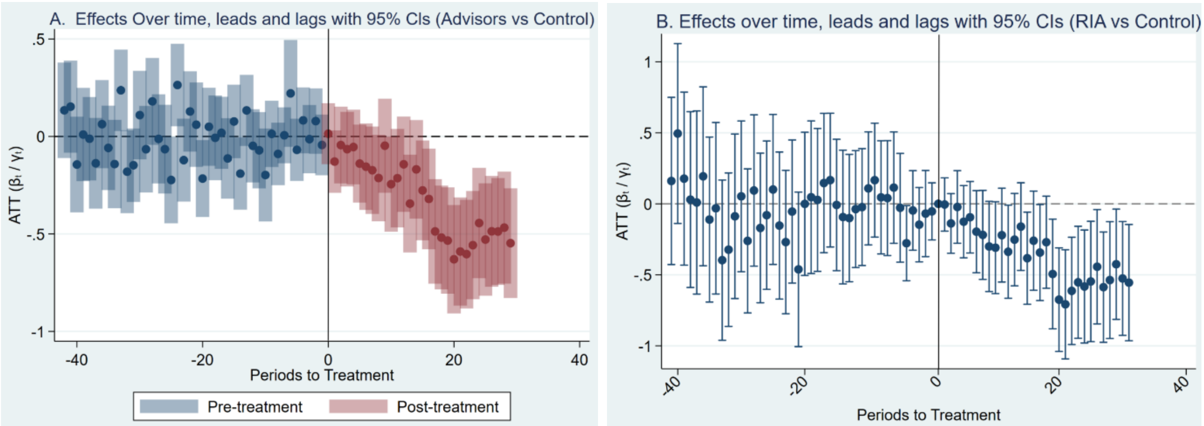

Fig. 10: Dynamic IPW-DiD

Each dot represents the point estimate for the corresponding lead or lag in Callaway and Sant'Anna (2021)'s IPW DiD model, estimated using Equations (4) to (8). Figure A compares professionals (both RIAs and unregistered) with control users, while Figure B compares RIAs with control users. Coefficients reflect differences in log message counts t months before or after the SEC policy change (ATT), relative to the baseline month ($t_0 = -1$). Error bars indicate 95% confidence intervals based on cluster-robust standard errors, adjusted for 69,314 user-level clusters.

Nearly all pre-policy ATTs (leads) are insignificant, supporting the no-anticipation assumption (treatment anticipation behavior). Post-policy, the reduction in professionals' posting activity attains statistical significance beginning in month 13 after the SEC event (July 2018) and intensifies through December 2020; See Fig. 10 (A). Specifically, RIAs' posting frequency declined by an average of 28.6% (calculated as exp(−0.337) − 1) over 2018-2020, with the reduction deepening to over 40% (exp(−0.520) − 1) by 2020; See Fig. 10 (B).

Altogether, the dynamic IPW-DiD model reinforces our principal finding: the SEC policy change exerted a significant chilling effect on professionals' social media use with particularly pronounced effects for RIAs. The SEC policy reduced RIAs' formal engagement with retail investors on social platforms and may have prompted broader adjustments to business practices. Two pieces of evidence support this interpretation: First, regulators have stated that "they seek to interpret rules flexibly to permit advisors' communications via social media," see FINRA, 2010. Second, the observed decline occurred despite RIAs' sustained high levels of content production and network influence during the four years preceding the SEC action, indicating prior efforts to expand their business via social media. This chilling effect may be driven by advisors' overestimation of the odds and costs of SEC punishment. Research in behavioral economics suggests that such overestimation can be attributed to cognitive biases, such as salience bias (Bordalo *et al.*, 2013) and local thinking (Gennaioli and Shleifer, 2010).[32]

5.3 Impact of SEC Restrictions: Robustness Test

In this section, we implement a **Trimmed DiD estimator** that explicitly enforces common-support trimming based on propensity scores. Specifically, treated and control units with estimated propensity scores outside the common support of the treated group are excluded to strengthen the validity of the parallel trends and no-anticipation assumptions. The remaining control and treated units are then equally weighted within their respective groups. This procedure refines the comparison group by removing poorly matched control and treated units, thereby supporting the identification assumptions. Unlike the IPW DiD estimator proposed by Callaway and Sant'Anna (2021), which retains all observations and reweights only control units, our approach imposes common-support restrictions prior to estimation for both control and treated users. The implementation proceeds in four steps:

1. **Propensity scores:** we estimate each user's treatment propensity score $\hat{p}_{iym} = P\left(treated_{iym} = 1 \,\middle|\, X_{iym}\right)$, using a logit model that predicts the likelihood of a user

[32] In salience bias, advisors place disproportionate weight on the costs of regulatory enforcement, overestimating them in relation to the benefits of using social media. In local thinking bias, judgments are influenced by what comes to mind based on prior experiences. Advisors apply their previous experiences with SEC enforcement (maybe unrelated to their use of social media) to their judgments about using social media following regulatory enforcements.

being a professional based on pre-SEC-policy network covariates X_{iym}. Where i indexes individual (user), y indexes years (2014- 2018), and m indexes months (1-12). Table A. 3 in Appendix provides a detailed report on the logit model.

2. **Average propensity score:** The average propensity score is then estimated for each user over the pre-treatment period, where T_i denotes the number of monthly observations for user i:

$$\hat{p}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} \hat{p}_{it}.$$

Fig. 11 displays the average propensity scores for treated and control units. Model adequacy is assessed based on overlap and covariate balance. As shown in the figure, the kernel densities of $\hat{p}_i$ for treated and control users exhibit substantial overlap over the [0,1] range, with little mass near 0 or 1 for treated units, indicating feasible adjustment.

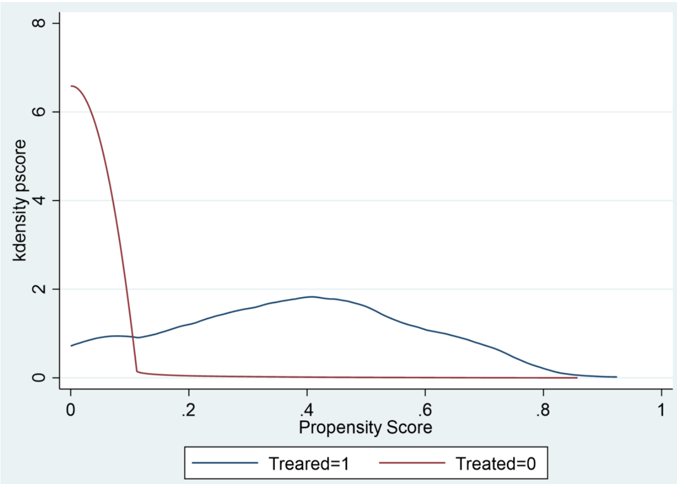

Fig. 11: kernel densities of $\hat{p}_i$ by treatment status

Figure shows kernel densities of average propensity scores ($\hat{p}_i$) for control and treated users during the pre-treatment period. Propensity scores are estimated based on covariates capturing users' network characteristics prior to the SEC event, including the number of Followers, Following, Watchlist Size, Ideas, Membership Days, and Likes per user. A detailed description of the logit model used to estimate $\hat{p}_i$ is provided in Table A. 3 of the Appendix.

3. **Common-support trimming:** Using a trial-and-error approach, we identify an appropriate common-support region for trimming to ensure that both the parallel trends and no-anticipation assumptions hold. Accordingly, we restrict the comparison sample to control users with $0.17 < \hat{p}_i \leq 1$ (see Fig. 11). Under this restriction, the no-anticipation test (based on Equation (4)) fails to reject the null of no effect in

anticipation of treatment $(F(42,874) = 1.21, p = 0.1773)$, and the pre-treatment parallel trends test likewise fails to reject the null of linear parallel trends $(F(1,874) = 2.40, p = 0.1218)$.

The linear-trends model used for the parallel-trends test augments the standard DiD specification in Equation (3) with two additional interaction terms. Let $d_{t,0} = 1(d_t = 0)$ denote the pre-SEC policy indicator and $d_{t,1} = 1(d_t = 1)$ denote the post-SEC policy indicator. Define $w_i = 1$ if user i belongs to the treated group (professionals) and $w_i = 0$ otherwise. The augmented model is specified as:

$$y_{it} = DiD_{it} + \zeta_1 w_i d_{t,0} t + \zeta_2 w_i d_{t,1} t + \varepsilon_{it}. \tag{9}$$

Where $DiD_{it} = \alpha + \beta D_{it} + \mu_i + \lambda_t$, as defined in Equation (3). In this specification, the coefficient ζ_1 captures differences in outcome slopes between the treated and control groups during the pre-SEC policy, while ζ_2 captures differences in slopes during the post-SEC policy. A value of $\zeta_1 = 0$ indicates that the linear trends in outcomes are parallel prior to treatment. We conduct a Wald test of $\zeta_1 = 0$ to evaluate whether the pretreatment linear trends are statistically indistinguishable between treated and control units $(F(1,874) = 2.40, p = 0.1218)$; see Fig. 12.

Notably, the overlap trimming ($0.17 < \hat{p}_i \leq 1$) substantially reduces the comparison sample: the number of control users decreases from 68,866 to 662 (a ~99% reduction), yielding a subset whose pre-policy network characteristics closely resemble those of professionals. The treated sample also contracts from 448 to 361 (a ~19% reduction), as only professionals within the overlap region are retained. This implies that the remaining control units are high-propensity users whose observable network characteristics closely mirror those of professionals; however, interpreting them as undisclosed professionals would be speculative without external validation.

4. **ATT estimation:** Finally, we estimate the ATT using Equations (3), and the results are presented in Table 12.[33]

[33] **Stata Command:** xtdidregress (logMessage) (did) if (0.17 < pscore) & (pscore <= 1), group(UserID) time(mdate) vce(cluster UserID).

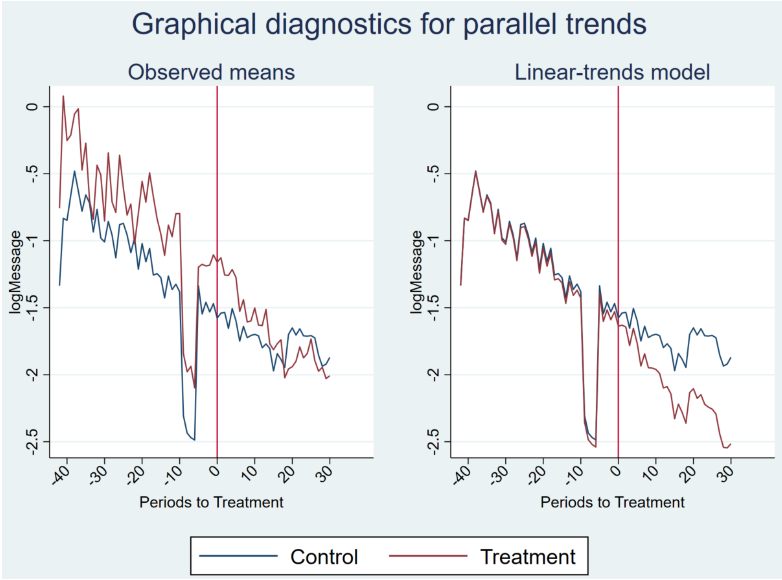

Fig. 12: Graphical Diagnostic of Parallel Trends using the Trimmed DiD.

This figure presents the graphical diagnostic of the parallel trends for the trimmed DiD estimator, estimated using Equation (9). The plot illustrates the pre- and post-SEC policy dynamics of posting activity for professionals and control users, showing parallel trends prior to the policy change and divergence thereafter, consistent with the estimated ATT effects.

Table 12: The ATT Using the Trimmed DiD

This table presents the results of the trimmed DiD ($0.17 < \hat{p}_i < 1$) for individual-level panel data, with time (months $t = 1, \ldots, T$, with $T = 84$ spanning January 2014–December 2020) and user fixed effects (control group: n = 662; treated group: n = 361), estimated using (3). The outcome variable is the logarithmically transformed number of monthly messages posted by a StockTwits user, defined as "log(monthly message volume + 0.5)." The SEC policy event is dated July 2018 (t = 54), when the Commission reaffirmed the applicability of federal securities laws to professionals' use of social media and announced five settlements. The treatment variable is a dummy variable equal to 1 for professionals (advisors) at time t and 0 otherwise. We restrict the sample to users with an established posting history on StockTwits prior to the SEC policy event. We construct a balanced panel by setting missing user–month outcomes to zero (no record ⇒ no relevant posts). Standard errors are clustered by user (1,023 clusters). The estimation sample contains 74,679 user–month observations.

	Coefficient	std. err.	t	P>|t|	[95% conf. interval]	
ATT	-0.258	0.127	-2.04	0.042	-0.506	-0.009

Consistent with prior results, Table 12 shows that the trimmed DiD estimator yields an ATT of −0.258, corresponding to a 22.7% decline in professionals' posting activity following the SEC restrictions ($\exp(-0.258) - 1$), statistically significant at the 5% level. This effect is approximately half the magnitude of the standard DiD estimate (46% decline, $\exp(-0.613) -$

1) reported in Model 4 of Table 11. The attenuation aligns with improved covariate balance and the restriction of the comparison sample to control and treated units closely matched to professionals' pre-SEC network characteristics, thereby reducing extrapolation from poorly matched observations.

Fig. 12 presents the graphical diagnostic of the parallel trends based on the trimmed DiD estimator, showing strong support for the parallel trends assumption and an intensifying post-SEC ATT through 2020. These results provide robust evidence of a persistent chilling effect of the SEC restrictions on the posting activity of professionals, who publicly disclosed their identities through official accounts.

6 Conclusion

This paper offers a novel analysis of the factors driving percolation and the spread of investment advice in finance-oriented social networks, with a focus on StockTwits. It further explores how professional investment advisors navigate social media and regulatory constraints to mitigate the risk of disintermediation posed by such platforms.

Initially, we focus on understanding the factors determining network influence. We observe that information dissemination on StockTwits often occurs off-structure, distinguishing well-connected members and effective influencers, unless well-connected members are identified as Official Accounts. We also discover that message content, specifically *Negative sentiment* and *Partisan nature*, plays a crucial role in driving influence within the network. Notably, less-connected members strategically exploit a negative tone and reference partisan news media to amplify their influence, contrasting with the strategies of well-connected members.

In examining the presence of investment advisors on StockTwits, our findings suggest that the network does not necessarily threaten established investment advice. Through its content curation mechanism, and the opportunity to be labeled as an Official Account, professionals are able to still control advice, even if the advisees may not easily be identifiable since they may no longer be structurally connected as followers in the network.

We find that RIAs leveraged StockTwits' Official Account feature to establish themselves as influencers within the network, generating viral cascades and driving information flows. Surprisingly, however, since 2018 RIAs have decreased their network activity. Our investigations attribute this reduction to regulatory enforcement by the SEC and FINRA. We conclude that financial disintermediation in investment advice is a consequence of regulatory overreach rather than the nature of social media itself.

We believe our findings offer meaningful insights into how regulatory interventions exerted a *chilling effect* on RIAs' formal engagement with retail investors on social platforms, potentially prompting broader adjustments in their business practices. The observed decline in network activity is striking, given RIAs' sustained content production and influence during the four years preceding the SEC action, reflecting their earlier efforts to expand business through social media.

Given our evidence that professionals—unlike other users—tend to avoid partisan or emotionally charged content, and prior research showing that their posts more accurately predict next-day stock returns than those of other StockTwits users (see Cookson et al., 2024), their withdrawal likely diminished the informativeness of socially sourced investment advice, thereby disadvantaging retail investors. Nonetheless, this interpretation remains conjectural and calls for external validation, offering a promising avenue for future research.

References

Bali, T.G., Hirshleifer, D., Peng, L., Tang, Y., 2021. Attention, social interaction, and investor attraction to lottery stocks. National Bureau of Economic Research.

Barabási, A.-L., 2009. Scale-free networks: a decade and beyond. Science, 325, 412-413, https://www.science.org/doi/10.1126/science.1173299.

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

Barthélemy, M., Barrat, A., Pastor-Satorras, R., Vespignani, A., 2004. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. Physical review letters, 92, 178701, https://doi.org/10.1103/PhysRevLett.92.178701.

Bild, D.R., Liu, Y., Dick, R.P., Mao, Z.M., Wallach, D.S., 2015. Aggregate characterization of user behavior in Twitter and analysis of the retweet graph. ACM Transactions on Internet Technology (TOIT), 15, 1-24, https://doi.org/10.1145/2700060.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and consumer choice. Journal of Political Economy, 121, 803-843.

Bradley, D., Hanousek Jr, J., Jame, R., Xiao, Z., 2024. Place your bets? The value of investment research on Reddit's Wallstreetbets. The Review of Financial Studies, 37, 1409-1459.

Broido, A.D., Clauset, A., 2019. Scale-free networks are rare. Nature Communications, 10, 1017, https://doi.org/10.1038/s41467-019-08746-5.

Bryzgalova, S., Pavlova, A., Sikorskaya, T., 2023. Retail trading in options and the rise of the big three wholesalers. The Journal of Finance, 78, 3465-3514.

Callaway, B., Sant'Anna, P.H., 2021. Difference-in-differences with multiple time periods. Journal of econometrics, 225, 200-230.

Cao, S., Jiang, W., Wang, J., Yang, B., 2024. From man vs. machine to man+ machine: The art and AI of stock analyses. Journal of Financial Economics, 160, 103910.

Chalmers, J., Reuter, J., 2020. Is conflicted investment advice better than no advice? Journal of Financial Economics, 138, 366-387.

Chen, S., Peng, L., Zhou, D., 2025. Wisdom or Whims? Decoding the Language of Retail Trading with Social Media and AI. Decoding the Language of Retail Trading with Social Media and AI (April 02, 2025).

Christoffersen, S.E., Evans, R., Musto, D.K., 2013. What do consumers' fund flows maximize? Evidence from their brokers' incentives. The Journal of Finance, 68, 201-235.

Clauset, A., Shalizi, C.R., Newman, M.E., 2009. Power-law distributions in empirical data. SIAM review, 51, 661-703, https://doi.org/10.1137/070710111.

Cookson, J.A., Engelberg, J.E., Mullins, W., 2023. Echo chambers. The Review of Financial Studies, 36, 450-500, https://doi.org/10.1093/rfs/hhac058.

Cookson, J.A., Lu, R., Mullins, W., Niessner, M., 2024. The social signal. Journal of Financial Economics, 158, 103870.

DeMarzo, P.M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. The Quarterly journal of economics, 118, 909-968.

Duguay, R., Lin, J., Thomas, J.K., Yin, L., 2024. Regulating via Social Media: Deterrence Effects of the SEC's Use of Twitter. Available at SSRN 3952904.

Eaton, G.W., Green, T.C., Roseman, B.S., Wu, Y., 2022. Retail trader sophistication and stock market quality: Evidence from brokerage outages. Journal of Financial Economics, 146, 502-528.

Gennaioli, N., Shleifer, A., 2010. What comes to mind. The Quarterly journal of economics, 125, 1399-1433.

Goel, S., Anderson, A., Hofman, J., Watts, D.J., 2016. The structural virality of online diffusion. Management Science, 62, 180-196, https://doi.org/10.1287/mnsc.2015.2158.

Goel, S., Watts, D.J., Goldstein, D.G., 2012. The structure of online diffusion networks. In: Proceedings of the 13th ACM conference on electronic commerce, pp. 623-638 https://doi.org/10.1145/2229012.2229058.

Golub, B., Jackson, M.O., 2010. Naive learning in social networks and the wisdom of crowds. American Economic Journal: Microeconomics, 2, 112-149.

Greig, F., Ramadorai, T., Rossi, A.G., Utkus, S.P., Walther, A., 2024. Algorithm aversion: Theory and evidence from robo-advice. Available at SSRN 4301514.

Hirshleifer, D., Peng, L., Wang, Q., 2025. News diffusion in social networks and stock market reactions. The Review of Financial Studies, 38, 883-937.

Kwak, H., Lee, C., Park, H., Moon, S., 2010. What is Twitter, a social network or a news media? In: Proceedings of the 19th international conference on World wide web, pp. 591-600 https://doi.org/10.1145/1772690.1772751.

Linnainmaa, J.T., Melzer, B.T., Previtero, A., 2021. The misguided beliefs of financial advisors. The Journal of Finance, 76, 587-621.

Myers, S.A., Sharma, A., Gupta, P., Lin, J., 2014. Information network or social network? The structure of the Twitter follow graph. In: Proceedings of the 23rd International Conference on World Wide Web, pp. 493-498 https://doi.org/10.1145/2567948.2576939.

Needham, M., Hodler, A.E., 2019. Graph algorithms: practical examples in Apache Spark and Neo4j. O'Reilly Media.

Pastor-Satorras, R., Vespignani, A., 2001. Epidemic spreading in scale-free networks. Physical review letters, 86, 3200, https://doi.org/10.1103/PhysRevLett.86.3200.

Vosoughi, S., Roy, D., Aral, S., 2018. The spread of true and false news online. science, 359, 1146-1151, https://www.science.org/doi/10.1126/science.aap9559.

Watts, D.J., 2002. A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99, 5766-5771, https://doi.org/10.1073/pnas.082090499.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of ‘small-world’networks. nature, 393, 440-442, https://doi.org/10.1038/30918.

Welch, I., 2022. The wisdom of the Robinhood crowd. The Journal of Finance, 77, 1489-1527.

Appendix

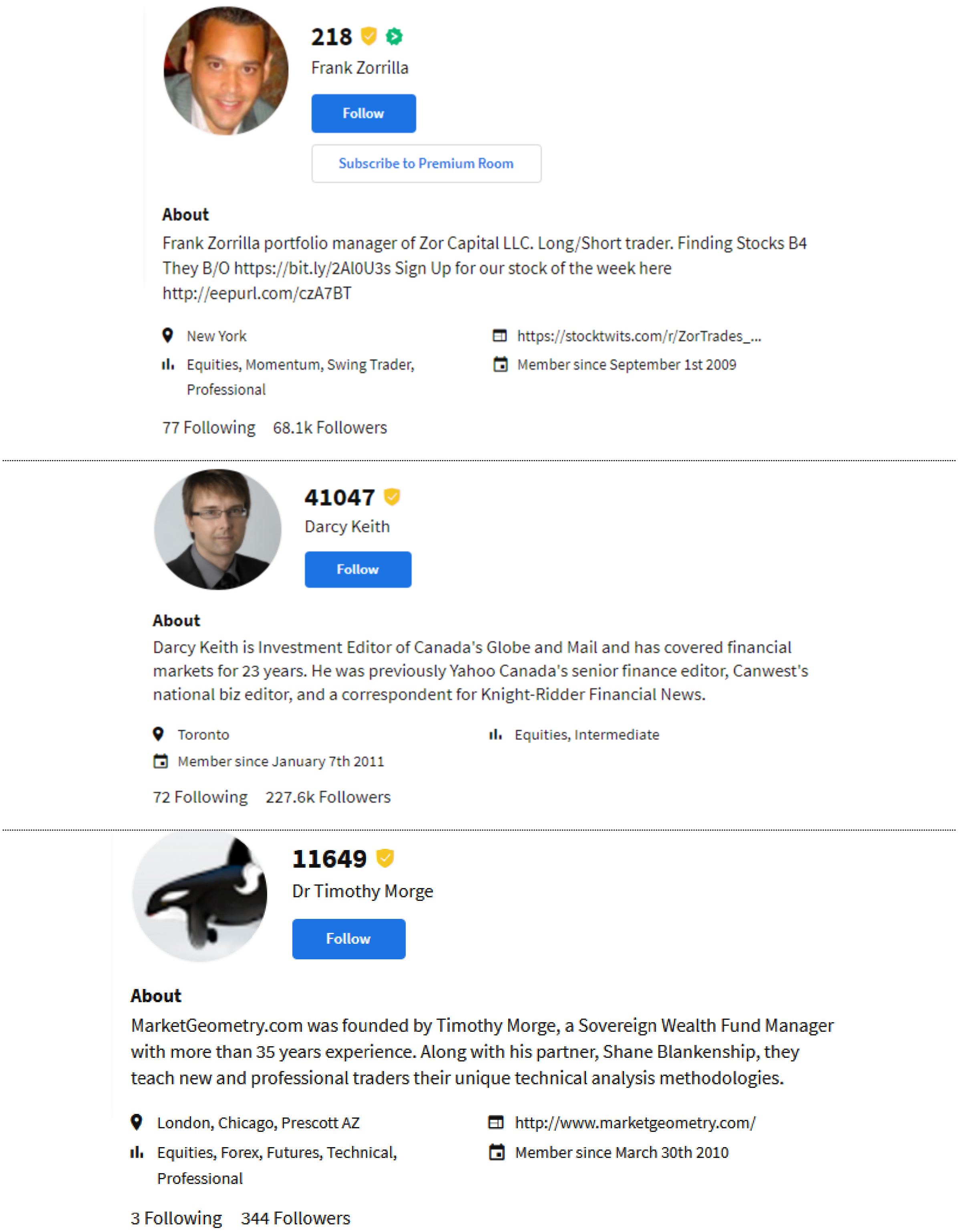

Fig. A. 1: Examples of Official Users on StockTwits

Table A. 1: Firms in our StockTwits Data, classified by Market Cap and GICS Sectors

Part 1)

	Communication Services	Consumer Discretionary	Consumer Staples	Energy
Large	Twitter Inc Netflix Inc Snap Inc	Starbucks Corp Tesla Inc Ford Motor Co	Sysco Corp Conagra Brands Inc General Mills Inc	Valero Energy Corp Concho Resources Inc Marathon Petroleum Corp
Medium	Yelp Inc Meredith Corp Cable One Inc	Under Armour Inc Grand Canyon Education Inc Vail Resorts Inc	Beyond Meat Inc Post Holdings Inc Flowers Foods Inc	National Oilwell Varco Inc Cimarex Energy Co Sunoco Lp
Small/Micro	Eventbrite Inc Gannett Co Inc Clear Channel Outdoor	Gopro Inc Xpresspa Group Inc Remark Holdings Inc	Andersons Inc Revlon Inc Weis Markets Inc	Camber Energy Inc Gevo Inc W&T Offshore Inc
Dow Jones	Verizon Communications Disney (Walt) Co	Home Depot Inc Nike Inc Mcdonald'S Corp	Walmart Inc Coca-Cola Co Procter & Gamble Walgreens Boots	Chevron Corp Exxon Mobil Corp

Part 2)

	Financials	Health Care	Industrials	Information Technology
Large	Globe Life Inc S&P Global Inc Lincoln National Corp	Alnylam Pharmaceuticals Inc Gilead Sciences Inc Dentsply Sirona Inc	Uber Technologies Inc American Airlines Group Inc Delta Air Lines Inc	Paypal Holdings Inc Micron Technology Inc Nvidia Corp
Medium	First Hawaiian Inc Independent Bank Corp/Ma Fs Kkr Capital Corp	Moderna Inc Hms Holdings Corp Bruker Corp	Virgin Galactic Holdings Inc Ryder System Inc Werner Enterprises Inc	Ncr Corp Littelfuse Inc Blackline Inc
Small/Micro	Ladder Capital Corp Eagle Bancorp Inc/Md Federal Agriculture Mtg Cp	Biocept Inc Ibio Inc Heat Biologics Inc	Plug Power Inc Fuelcell Energy Inc Hawaiian Holdings Inc	Vislink Technologies Inc Inpixon Fitbit Inc
Dow Jones	American Express Co Goldman Sachs Group Inc Jpmorgan Chase & Co Travelers Cos Inc	Johnson & Johnson Merck & Co Pfizer Inc Unitedhealth Group Inc	3M Co Boeing Co Caterpillar Inc Raytheon Technologies Corp	Visa Inc Apple Inc Cisco Systems Inc Intl Business Machines Corp Intel Corp Microsoft Corp

Part 3)

	Materials	Real Estate	Utilities
Large	Eastman Chemical Co Ball Corp Vulcan Materials Co	Vornado Realty Trust Equity Residential Sba Communications Corp	Essential Utilities Inc Centerpoint Energy Inc Vistra Corp
Medium	Sonoco Products Co Sensient Technologies Corp Innospec Inc	Healthcare Trust Of America Kilroy Realty Corp Potlatchdeltic Corp	South Jersey Industries Inc Idacorp Inc Northwest Natural Hldng Co
Small/Micro	Valhi Inc Worthington Industries Mesabi Trust	Washington Prime Group Inc Universal Health Rlty Income Hersha Hospitality Trust	Unitil Corp York Water Co Middlesex Water Co
Dow Jones	Dow Inc		

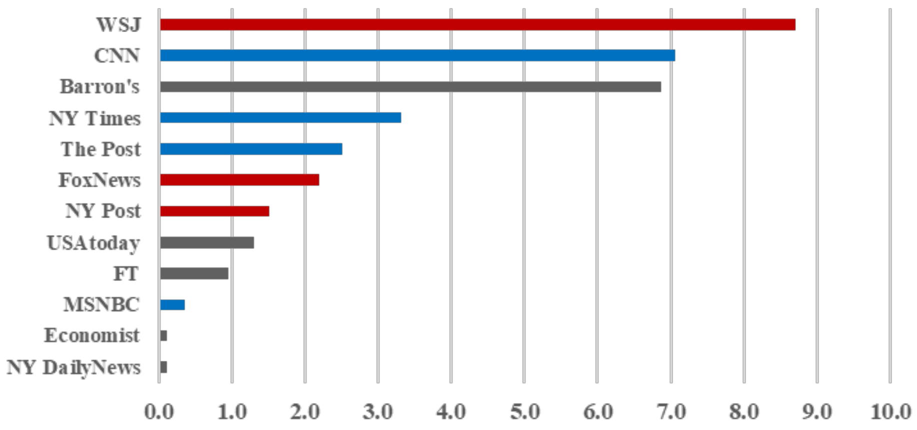

Fig. A. 2: Distribution of Number of Cascades Across News Outlets.

The figure displays the distribution of the number (X1,000) of cascades in StockTwits across news outlets based on the news outlet that the originating message referred to, stratified by political polarization. Red, gray, and blue bars show "right-center partisan," "least-biased media," and "left-center partisan," respectively.

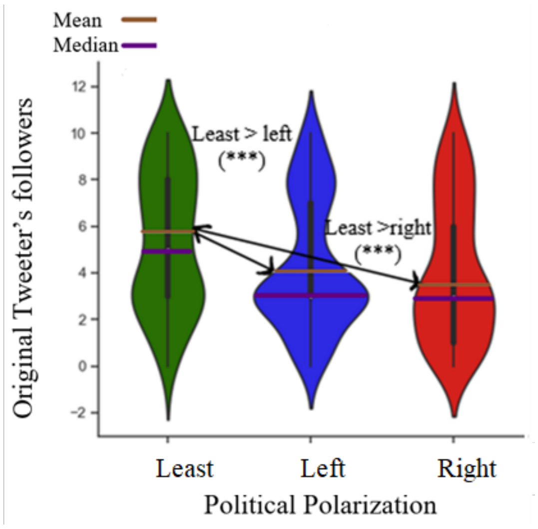

Fig. A. 3: Follower Count of Original Tweeters of Bearish News Story Cascades across News Outlets

Figure displays three violin plots illustrating the follower count distributions of original tweeters. The analysis is limited to posts, which are labeled as Bearish (negative) and referencing twelve different news outlets. Y-axis represents the follower count of original tweeters, while the X-axis categorizes tweets based on the partisan nature (least-biased, left, and right) of the news outlets mentioned in posts. In each violin plot, horizontal lines display the mean (brown) and median (purple) of the distribution. The thick black bar in the center presents the interquartile range, and the thin black line represents the rest of the distribution, except for points that are deemed to be outliers. On each side of the black line is an estimate of the probability density based on kernel estimation. The distribution functions are compared using KS-statistics; stars (***) denote statistical significance at $p < 0.001$.

Table A. 2: Examination of Social Attention in StockTwits

This table presents the outcomes of a random-effects multinomial logistic regression analysis (panel variables: firm and user) that estimates the probability of Social Attention for a post on StockTwits—specifically, the number of likes received by the corresponding cascade, using the following equation (Number observations = 1,585,082, χ^2 =240939.83, Log likelihood = -1125533.4, and Pseudo R^2 = 0.0967):

$$\log\left(\frac{prob\ (Y=1)}{prob\ (Y=0)}\right) = \alpha_1 + \alpha_2 + \beta_1\ Potential\ influencer + \beta_2\ Official\ user + \beta_3\ Sentiment + \beta_4\ Partisan\ nature + User + Date + Firm + \varepsilon_0; \quad i = 2\ and\ 3.$$

The outcome variable (Y) is categorical, comprising Category 1 (no likes received by the cascade), Category 2 (0 < likes received by the cascade < 4), and Category 3 (4 ≤ likes received by the cascade). Specifically, in each regression analysis, Category 1 (small cascades) serves as Y=0 (base variable), and one of the other categories (either 2 or 3) represents the measurement of Y =1. α_1 and α_2 indicate random effects per message sender and per firm mentioned in the message, respectively. The model incorporates predictors including the characteristics of the original tweeter, partisan nature of the content, and the self-declared sentiment label of the post. Original tweeter characteristics are measured through "Potential influencer," "Official user," "Investment advisor," and "RIA." "Potential Influencer" is a dummy variable equal to 1 if the original tweeter does have more than 1000 followers. "Official users" is a dummy variable equal to 1 if the original tweeter holds an "Official Account." "Investment advisor" is a dummy variable equals to 1 if the original twitter holds an official account and is a professional in the investment field associated with investment advice companies or websites providing investment tools. "RIA" is a dummy variable set to 1 if the original tweeter holds an official account and is a "Registered Investment Advisor." "Sentiment" is a categorical variable equal to 1 and 2 if the original tweeter has labeled his/her post as "Bullish" and "Bearish," respectively. "Sentiment" equal to 0 if the original tweeter has not labeled his/her post. "Partisan nature" is a dummy variable equal to 1 if the original tweeter's post linking to a news article published by media deemed partisan (WSJ, NY Times, The Post, NY Post, CNN, FoxNews, and MSNBC). To mitigate the impact of user-specific, firm-specific, and seasonal variations on the estimation of the model, we control user, firm, and date fixed effects.

Y=1	Predictors	Odds Ratio	Std. Err.	Z	p-value	95% Confid. Interval	
Like Category 2 (0 < Likes< 4)	Potential Influencer	0.545	0.002	-133.87	<0.001	0.540	0.550
	Official user	1.130	0.009	14.58	<0.001	1.111	1.149
	Investment advisor	1.100	0.010	10.36	<0.001	1.081	1.120
	RIA	5.132	0.105	79.77	<0.001	4.930	5.342
	Partisan nature	1.868	0.037	31.63	<0.001	1.797	1.941
	Negative sentiment	3.448	0.019	222.24	<0.001	3.411	3.486
	Positive sentiment	1.920	0.019	64.3	<0.001	1.883	1.959
	Constant	0.361	0.002	-158.37	<0.001	0.356	0.365
Like Category 3 (4 ≤ Likes)	Potential Influencer	0.387	0.004	-98.35	<0.001	0.380	0.395
	Official user	1.363	0.032	13.13	<0.001	1.301	1.428
	Investment advisor	0.848	0.023	-6.14	<0.001	0.805	0.894
	RIA	18.121	0.591	88.83	<0.001	16.999	19.317
	Partisan nature	2.416	0.066	32.51	<0.001	2.291	2.548
	Negative sentiment	10.225	0.088	269.62	<0.001	10.053	10.399
	Positive sentiment	1.726	0.037	25.28	<0.001	1.654	1.801
	Constant	0.017	0.000	-263.1	<0.001	0.017	0.018

IPW DiD Model

We follow the IPW DiD of Callaway and Sant'Anna (2021) for identification, estimation, and inference of the SEC restrictions on professionals, using all treated and non-treated (control) units. Our setup includes the following steps:

1. **Estimate Propensity Scores:** The propensity score is the probability of receiving treatment, for each individual (user). This is estimated conditional on baseline covariates. For each user i, we estimate the (pre-treatment) propensity score:

$$p_{it} = P\,(G_i = g^* \mid X_{it}),\ t \leq g^* - 1. \qquad \text{(A1)}$$

We then use pre-treatment-period average propensity,

$$p_i(X) \equiv \frac{1}{g* - 1} \sum_{i=1}^{g* - 1} p_{it}. \qquad \text{(A2)}$$

Users i = 1, …, 69,314 (control n = 68,866; treated n = 448); months t = 1, …, T with T = 84 spanning January 2014–December 2020. All treated users (professionals: RIA and UnRIA) first receive treatment at the common date g^* = 54 (July 2018); never-treated users serve as controls. Where $G_i = g*$ for professionals and $G_i = 0$ for controls.

All covariates (X) enter the regression as month–year specific **decile ranks (1–10)**. "Monthly" covariates are measured within month m; "Cumulative" covariates are measured from the user's join date through the end of month m. Covariates include Followers (monthly avg.), Following (monthly avg.), Watchlist size (monthly avg.), Ideas (cumulative posts), UserLikes (cumulative likes), and Membership days (cumulative). Refer to Table A. 3 for a detailed description of the logit model used to calculate propensity scores.

2. **Stabilized IPW weights:** For post-period $t \geq g^*$, define the (normalized) *control weight*:

$$\omega_i(X) = \frac{\frac{p_i(X)}{1-P_i(X)}}{E\left[\frac{p_j(X)}{1-p_j(X)} \mid G=0\right]}. \qquad \text{(A3)}$$

and set $\omega_i(X) \equiv 1$ For professionals ($G_i = g^*$). These weights create a covariate-balanced pseudo-population in which controls mirror the treated cohort.

3. **Estimate ATT by weighted DiD:** estimate the average treatment effect on the treated (ATT) via a weighted DiD.

$$\text{ATT}(g, t) = E[Y_{i,t} - Y_{i, g^*-1} \mid G = g^*] - E_{\omega}[Y_{i,t} - Y_{i, g^*-1} \mid G = 0]. \quad \text{(A4)}$$

Where, $E_{\omega}[\cdot \mid Gi = 0]$ denotes the expectation under weights $\omega_i(X)$, and the ATT indicates the IPW DiD estimation of coefficient β in the standard TWFE model:

$$y_{it} = \beta D_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \; i = 1, \dots, N; \;\; t = 1, \dots, T. \quad \text{(A5)}$$

This model has a full set of time effects λ_t, a full set of group effects, D_{it} is a variable that is 1 if an individual is treated at time t and is 0 otherwise, and ε_{it} is an error term. The outcome variable of interest y_{it} is defined as the logarithm of the number of messages (logMessage) posted by user i in period t, aggregated at the monthly (year–month) level from January 2014 to December 2020. We construct a balanced panel by coding missing values of y_{it} as zero, under the assumption that the absence of a record on a given day indicates that the user did not post about the relevant stocks.

Table A. 3: Logit Model for Calculating Propensity Scores

This table reports estimates from a logistic regression of the probability of receiving the treatment (the propensity score) on users' pre-treatment network characteristics (covariates), estimated using the unbalanced panel. The panel comprises 294,754 observations. The model yields a log pseudolikelihood of −17,015.3, a Wald $\chi^2(48)$ statistic of 768.48, and a pseudo-R^2 of 0.5545. Standard errors are clustered at the user level (69,314 clusters). The estimated specification is given by:

$$\log\left(\frac{p_{iym}}{1-p_{iym}}\right) = \beta_0 + \beta' X_{iym} + \lambda_t + \varepsilon_{iym}\,, \quad p_{iym} = P\left(treated_{iym} = 1 \mid X_{iym}\right).$$

where i indexes individuals (User IDs), y indexes years (2014–2018), and m indexes months (1–12). λ_t represents time effects, captured by year–month dummies, and ε_{iym} denotes individual-specific error terms. X_{iym} is a vector of user i's observable characteristics in month m of year y (covariates). The variable treated$_{iym}$ is a binary indicator equal to 1 if the user is either an RIA or an UnRIA, and 0 otherwise. All covariates enter the regression as month–year specific decile ranks (1–10). "Monthly" covariates are measured within month m, whereas "cumulative" covariates are measured from the user's join date through the end of month m. The covariates are defined as follows:[34]

Followers (monthly average): Average number of followers of user i during month m, decile-ranked within month. **Following (monthly average):** Average number of accounts followed by user i during month m, decile-ranked within month. **Watchlist size (monthly average):** Average number of securities on user i's watchlist during month m, decile-ranked within month. **Ideas (cumulative posts):** Total number of messages posted by user i from their join date through the end of month m, decile-ranked within month. **UserLikes (cumulative likes):** Total number of likes received by user i from their join date through the end of month m in year y, decile-ranked within month**. Membership days (cumulative):** Number of days since user i's join date through the end of month m in year y, decile-ranked within month.

Note: "Decile-ranked within month" indicates that ranks are computed cross-sectionally among all users observed in the same month, with one decile indicator omitted for identification.

	Odds Ratio	Robust Std. Err.	Z	p-value	95% Confid. Interval	
Followers	16.003	2.763	16.06	< 0.0001	11.408	22.448
Following	1.174	0.028	6.74	< 0.0001	1.120	1.229
Membership Days	1.316	0.043	8.33	< 0.0001	1.233	1.403
Ideas	0.864	0.036	-3.47	0.001	0.795	0.938
Watchlist	0.887	0.022	-4.80	< 0.0001	0.845	0.932
UserLike	0.910	0.025	-3.40	0.001	0.862	0.961

[34] **Stata Command:** logit treated Userfollowers_decile Userfollowing_decile userMembershipDays_decile Userideas_decile watchlist_decile Userlike_decile i.mdate, or vce (cluster UserID).