Encouraging Broker and Adviser Background Checks: A Randomized Study on Twitter*

K. Jeremy Ko† Sai Rao†† Parth Venkat††

November 29, 2018

Abstract:

We conducted a randomized online study of messaging on twitter to encourage retail investors to check the background of their investment professional. Our study tested different tweets attempting to motivate investors by emphasizing aspirations, mistrust, and loss aversion. We found that our motivational tweets did not perform better than our control informational tweet. In contrast, there appears to be evidence that shorter, simpler calls to action worked best in promoting views, engagements, and clicks on the SEC website. There is also some evidence of greater (lower) engagement at the beginning (end) of the week. Finally, there was only limited attrition in views and engagements over time with high turnover in accounts engaging tweets over the course of the study.

Keywords: randomized control trial, social media, brokers, investment advisers, behavioral finance

JEL Codes: D14, D18, G02, G24

---

* The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This article expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the staff. We thank Owen Donley, Anna Huntley, DaVeda Johnson, Nathasha Lim, Brian Mulford, Claire O'Sullivan, the SEC Office of Investor Education and Advocacy, and the SEC Office of Public Affairs for their invaluable support on this project. We also thank Vibha Puri, Leanne Robinson, and Jack Tu for their research assistance as well as seminar participants at the Securities and Exchange Commission and the General Services Administration for their helpful comments. All errors are our own.

† Corresponding author: Division of Economic and Risk Analysis, U.S. Securities and Exchange Commission. 100 F St NE, Washington, DC 20549, kok@sec.gov, Phone: 202-551-7895.

†† Division of Economic and Risk Analysis, U.S. Securities and Exchange Commission.

Introduction

This study tests the effect of different types of messaging on encouraging individual or "retail" investors to check the background of their investment professionals (e,g., their individual adviser or broker providing investment advice) online. It has been documented that only a small fraction of US investors have ever used or are even aware of regulatory websites (i.e., FINRA's BrokerCheck and the SEC's Investment Adviser Public Disclosure websites), which enable background checks of registered investment professionals. For example, a 2015 survey of US retail investors conducted by the FINRA Investor Education Foundation found that only 7% of investors have ever used BrokerCheck, while only 16% were aware of it.[1]

Background checks are an important component of due diligence on investment professionals. A number of studies have shown that prior misconduct, legal infractions, and conflicts of interest can predict future misconduct by investment professionals. In addition, it is generally known that a large proportion of prosecuted investment fraud and misconduct is allegedly conducted by unregistered intermediaries. Therefore, background checks can potentially deter fraud and misconduct primarily by directing them to registered intermediaries and secondarily by diverting them from professionals with a misconduct record.

We deployed our messaging through tweets broadcast by twitter accounts managed by staff of the Securities and Exchange Commission. A growing number of studies document that tweets can have an impact on pricing and trading volume in financial markets.[2] Therefore, tweets are a potentially effective means for reaching and influencing the investing public. Our study tested five different tweets: one was a simple informational tweet (i.e., our experimental control) while the other four attempted to motivate investors by emphasizing psychological drivers such as aspirations, mistrust, and loss aversion (i.e., our experimental treatments). All tweets included a link to the SEC's investor education page (investor.gov), which features a prominent link to the SEC's Investment Advisor Public Database (IAPD) website with information about registered investment advisers. We hypothesize that these motivational messages will outperform the control informational tweet in terms of encouraging retail investors to conduct background checks. A number of studies find effects from such behavioral messaging primarily from the emphasis of

[1] The Dodd-Frank Section 917 Financial Literacy Study also found that 19% of investment advisory clients had ever used he IAPD website while 28% were aware of it.

[2] See, e.g., Gu and Chen (2018), Bartov, Faurel, and Mohanram (2018), Gu and Kurov (2018), Oliveira, Cortez, and Areal (2017), Mao, Counts, and Bollen (2015), Mao, Wang, Wei, and Liu (2012), and Bollen, Mao, and Zeng (2010).

social norms or peer comparisons.[3] However, other research shows limited effect from behavioral messaging.[4]

Standard twitter accounts do not allow for randomized A/B testing of tweets whereby the web server randomly would assign one among several alternative tweets to a particular user. A/B testing has become a standard method for testing design and messaging for websites and applications. Therefore, we attempted to randomize our testing by sending out one tweet each weekday in random order at a fixed time. We repeated this process for five weeks with the constraint that each tweet should be sent once each weekday. For example, tweet 1 appeared once on a Monday, once on a Tuesday, etc. This balanced design allowed us to simply compare averages of our dependent variables for each tweet to without controlling for week or day-of-the week. We could similarly compare averages for each week and for each the day-of-week. Given the duration of our study, we also tested the hypothesis of users suffering from "message fatigue" over the course of five weeks. In other words, there may be a significant loss in tweet engagements (i.e., retweets, replies, likes, link clicks, etc.) over this time.

In terms of our findings, we did not find evidence that our motivational tweets performed better than our control informational tweet. In fact, we found some evidence that our control tweet actually performed *better* than the treatment tweets. In addition, our tweet referencing losses and misconduct by investment professionals appears to have performed worse than the others. One possible ex-post interpretation of our findings is that shorter, simpler calls to action may have worked best in promoting views and engagements with the tweet. In addition, there was some evidence of attrition in views and engagements over course of our study. However, there was still significant engagement toward the end of the study with substantial turnover in accounts engaging with our tweets. Therefore, our audience did not exhibit dramatic "message fatigue" over the five weeks of our study.

A few other findings also emerged from our study. First, there was also some evidence of greater (lower) engagement at the beginning (end) of the week. In addition, a significant proportion of replies to our study tweets were negative in nature and appear to have come from spam-oriented accounts. However, the correlation between replies and link clicks was still positive, consistent with the view that "there's no such thing as bad publicity" on twitter.

---

[3] See, e.g., Schultz (1999), Cialdini and Goldstein (2004), Alcott (2009), Gerber, Green, and Larimer (2008), and Beshears, Choi, Laibson, Madrian, and Milkman (2015), and Hallsworth, List, Metcalfe, and Vlaev (2017).
[4] See, e.g., Chong, Karlan, Shapiro, and Zinman (World Bank Economic Review, 2015) and Bhargava and Manoli (AER 2015) on the ineffectiveness of messaging emphasizing social benefits and overcoming stigma, respectively. SBST (2015) reports limited impact from peer comparisons in US tax payments.

The rest of this paper proceeds as follows. Section 2 discusses the design of our randomized study and lays out our hypotheses. Section 3 discusses our findings, while section 4 concludes.

I. Research Design

We tested five different tweets intended to induce retail investors to conduct background checks by emphasizing specific motivations or psychological drivers. All tweets directed twitter users to the website of the SEC's Office of Investor Education and Advocacy (OIEA): investor.gov. The top of this site features a prominent link to the SEC's Investment Adviser Public Disclosure (IAPD) website, where users can retrieve mandatory disclosures for all registered investment advisers in the US.[5] We did not provide a link to FINRA's BrokerCheck as IAPD cross references this database. The list of tweets from our study is given below:

1. **Control:**
   Did you know? You can check the background of your investment professional here: investor.gov
2. **Aspiration:**
   Smart investors do their homework and check their investment professional's background. Check on yours: investor.gov
3. **Mistrust 1:**
   Some studies suggest that 1 in 14 investment professionals have a record of misconduct! Check on yours: investor.gov
4. **Mistrust 2:**
   Fraud is often conducted by unregistered investment professionals! Check that yours is registered here: investor.gov
5. **Loss Aversion:**
   Avoid losing money to misconduct by an unlicensed investment professional! Check that yours is licensed here: investor.gov

Tweet 1 is our control tweet which provides information about the link without appealing to a particular emotional or psychological motive/driver. Tweet 2 appeals to aspirational motives on the part of the investor desiring to identify themselves as "smart" investors. Such aspirational messaging is commonly used in advertising and has been used in messaging by the SEC's OIEA.[6] Tweets 3 and 4 appeal to mistrust or suspicion of investment professionals by highlighting the possibility of their misconduct. Survey-based studies have shown that trust is associated with investment and use of financial advice.[7] Therefore, we hypothesize that inducing mistrust will motivate investors to conduct due diligence and consider the option of terminating an advisory relationship. Tweet 5 appeals to loss

---

[5] We were requested by OIEA to have our tweets direct users to investor.gov rather than the IAPD website, since they are trying to market the former as the "go-to" site for retail investors.
[6] See Dimofte, Goodstein, and Brumbaugh (2014) for a study of aspirational messaging in advertising.
[7] See, e.g., Burke and Hung (2015).

aversion by adding the prospect of loss to fraud conducted by unlicensed investment professionals. Loss aversion refers to the propensity to avoid losses as a result of heightened disutility from losses relative to utility from gains, first established in studies by Tversky and Kahneman (1992) and studied in messaging by Hallsworth, et al. (2014).

We would ideally like to compare the number of background checks caused by these five tweets. Unfortunately, we are unable to observe the number of background checks on IAPD caused by these tweets. The closest observable proxy is the number of clicks on the tweet link. However, one strength of our study relative to other messaging studies is that we can draw a strong causal connection between our tweet and the observed behavior. In particular, we can observe the number of link clicks directly on the particular message. We also analyze two other dependent variables: 1.) engagements, i.e., the number of times users interacted with a particular tweet equal to the sum of link clicks, retweets, replies, likes, profile and detail expands, and follows; 2.) impressions or the number of potential views (i.e., times the tweet was sent to a user's timeline).

Our principal hypothesis is as follows:

Hypothesis 1: Tweets 2 through 5 should draw more link clicks than tweet 1.

In particular, tweets 2 through 5 attempt to motivate investors based on psychological drivers, whereas tweet 1 provides factual information devoid of such triggers. A secondary hypothesis is that our twitter audience may suffer from "message fatigue" after seeing similar tweets over the five weeks of each phase of the study. In particular, there may be a significant loss in tweet engagements (i.e., retweets, replies, likes, link clicks, etc.) over this time. There may also be a loss in impressions because the twitter algorithm tends to prioritize tweets with higher engagements.[8] Therefore, we have the following hypothesis:

Hypothesis 2: Tweets during week 1 should draw the highest average impressions and engagements, while tweets during week 5 should draw the lowest.

In our analysis, we compare averages for given tweets and weeks to test these hypotheses. We also examine averages for each day of the week to study differences in responses across days.

[8] See, e.g.: http://www.slate.com/articles/technology/cover_story/2017/03/twitter_s_timeline_algorithm_and_its_effect_on_us_explained.html.

a. Timing

Tweets were automatically posted on the SEC_DERA account each weekday at 2 PM during our study. They were then manually retweeted around 2:30 PM the same day by either the SEC_News account in phase 1 or the SEC_Investor_Ed account in phase 2. We then manually collected data on link clicks, engagements, and impressions at roughly 2 PM the next day.

Table 1 shows the order of the tweets sent during both phases of the study. Phase 1 took place during the five weeks from Monday, January 22nd until Friday, February 23rd, 2018. Tweets during this phase were retweeted by the general SEC_News account with an audience of roughly 200 thousand followers at the time. As mentioned previously, each tweet (1-5) was sent once per week in random order with the constraint that each tweet should be sent once each weekday, e.g., tweet 1 appeared once on a Monday, once on a Tuesday, etc. This balanced design allows us to simply compare averages of our dependent variables for each tweet to assess which one worked best or worst without controlling for week or day-of-the week. We can similarly compare averages for each week and for each the day-of-week.

Phase 2 of the study took place during the five weeks from Monday, April 2nd until Friday, May 4th, 2018. Tweets during this phase were retweeted through the account managed by SEC's OIEA, SEC_Investor_Ed, with an audience of roughly 60 thousand at the time. We experienced a retweeting error on Friday of the third week of phase 2 (Friday, April 20th) when the retweet was not deployed by the SEC_Investor_Ed account. For this reason, we discarded the data from week 3 of phase 2. Instead, we added an additional week of the study from Monday, June 11th until Friday, June 15th with the same tweet order as week 3. In addition, three of our tweets from phase 2 were deployed on the same day as tweets from the SEC_Investor_Ed account also encouraging followers to check the background of investment professionals on investor.gov. There was no substantial difference in our dependent variables for these three tweets versus the others in our study.[9]

There could, in principle, be differences in engagements and impressions over time because of increases or decreases in followership for our accounts. Indeed, followership increases uniformly for all accounts over the course of our study. Therefore, we normalize our dependent variables by total followership for the accounts that disseminate our study tweets.

[9] In particular, the mean number of link clicks, engagements, and impressions for these three "paired" tweets were 4, 14.6, and 1808, respectively. The mean number of link clicks, engagements, and impressions for the twenty-two "non-paired" tweets were 3.95, 15.8, and 1830, respectively.

II. Results

Figures 1 and 2 show graphs of link clicks, engagements, and impressions per thousand followers for weeks 1-5 of phase 1 and 2 of our study. The amount of activity from followers of SEC_News in phase 1 was roughly equivalent to the amount from followers of SEC_Investor_Ed in phase 2. We next discuss which tweets performed well or poorly on average. We also discuss activity by week and by day-of-the-week.

We use a bootstrapping methodology to assess statistical significance associated with differences in these averages. In particular, we generated pseudo-samples for each phase by drawing random observations from our sample without replacement. In other words, each pseudo-sample was a random reordering of observations. We then generate a distribution of each statistic across pseudo-samples from which we compute p-values for the estimate from the actual sample.

a. Findings by Tweet

Table 2 shows the average differences in our dependent variables between tweets 2 through 5 and tweet 1. We see that tweets 2 through 5 actually performed worse, on average, than tweet 1 during phase 1 of the study. This difference for tweet 4 is statistically significant at the 10% level for tweet 4, while the differences for tweets 2, 3, and 5 are significant at the 5% level. For phase 2, only tweet 5 did worse than tweet 1 at a statistically significant level of 5%. Therefore, our findings appear to be inconsistent with hypothesis 1 that behavioral messaging was effective in drawing more link clicks than a plain informational tweet.

We further explore our findings by reporting the tweets with the minimum and maximum link clicks, engagements, and impressions for both phases of our study in table 3. During phase 1, only tweet 1 was statistically significant in accruing the maximum link clicks ($p < 10\%$). During phase 2, only tweet 5 was statistically significant in generating the minimum link clicks ($p < 5\%$).

There was consistency, however, between phases 1 and 2 in which tweets experienced the most and fewest interactions and views when not accounting for statistical significance. In particular, both tweets 1 and 4 earned the highest interactions and views in both phases, while tweets 2 and 5 earned the fewest. This pattern appears consistent with the effectiveness of shorter, simpler messages. In particular, tweet 1 had the fewest number of characters while tweet 5 had the highest. It was also pithier in terms of nouns (not counting pronouns), cueing users with the terms “background” and “professional”. Tweet 4 was also pithy, cueing users with the terms “fraud” and “professionals”. In contrast, tweet 2 used four nouns (i.e., “investors”, “homework”, “professionals”, and “background”) while tweet 5 used three (i.e., “money”, “misconduct”, and “professional”).

Other studies suggest that shorter, simpler text may enhance information processing and attention.[10] In this study, our finding is complicated by the fact that certain tweets may have taken up more space on certain screens than others. For example, tweet 1 was three lines long on a four-inch mobile screen, while the other tweets were four lines long. Consequently, the link for tweet 1 may have been less likely cut off if it were to appear at the bottom of a mobile screen. Therefore, it remains for future study to test the effectiveness of simpler and shorter online messages more rigorously.

b. Findings by Day-of-Week

Table 4 shows which days of the week accrued the maximum and minimum link clicks, engagements, and impressions for phase 1 and 2 of our study, respectively. During phase 1, Monday was statistically significant in accruing the maximum engagements ($p < 10\%$), while Friday was statistically significant in accruing the minimum impressions ($p < 1\%$). During phase 2, only Monday was statistically significant in generating the maximum impressions ($p < 5\%$).

Therefore, there was some evidence of greater interactions and views with our tweets earlier than later in the week. However, the findings in this direction were somewhat muddy. For example, Thursday generated the maximum link clicks and engagements in phase 1 of our study, while generating the minimum link clicks and engagements in phase 2. These findings were not statistically significant, however.

c. Findings by Week

Table 5 shows which weeks of our study accrued the maximum and minimum link clicks, engagements, and impressions for phase 1 and 2 of our study, respectively. During phase 1, only week 1 was statistically significant in accruing the maximum link clicks ($p < 5\%$). During phase 2, week 2 was statistically significant in generating the maximum engagements ($p < 5\%$), while week 5 was statistically significant in generating the minimum engagements ($p < 5\%$).

However, week 2 received the minimum impressions in phase 1 of the study. In addition, week 1 received the minimum link clicks in phase 2 of the study. Although neither statistic was significant, these findings show that the attrition in impressions and engagements over the five weeks of both phases was

[10] For example, several studies find that shorter and simpler text in corporate disclosures enhance processing of information (e.g., Hwang and Kim, 2017; Loughran and MacDonald, 2014; Lawrence, 2013; Lehavy, Li, and Markley, 2011; You and Zhang, 2009).

not uniform over time. In addition, there was no dramatic loss in activity at any point. Therefore, although there was some evidence in support of hypothesis 2, this evidence was far from overwhelming.

We can see from figure 1 that there was slight attrition in activity from week 1 to week 5 in phase 1. This attrition amounted to a loss of about 10 engagements of roughly 40-60 per tweet in week 5 relative to week 1 and 1000 impressions of roughly 5000-6000 per tweet. We can also see from figure 1 that there was slight attrition in engagements, but not a significant loss in impressions from week 1 to 5 in phase 2. The attrition in engagements amounted to roughly 5 of 10-25 from week 1 to 5. One reason we may see less attrition in phase 2 is that the SEC_Investor_Ed deploys tweets encouraging background checks on a regular basis. As a result, we may be observing responses from an audience already saturated with this particular message.

Among the accounts engaging with the tweets from our study, we see a high degree of turnover. Table 6 shows the amount of repeat interactions (in terms of likes, replies, and retweets) by accounts. The median number of interactions per account among these interacting accounts was 1 while the mean was slightly higher at 1.5 combining both phases of the study. In addition, the percent of users interacting with our study tweets only once was 70% (142 of 202). Therefore, there was a high degree of refreshment or turnover in the accounts interacting with these tweets.

We conclude that there may be legitimate concerns about "message fatigue" in extended repeat twitter campaigns. However, there appears to be a significant pool of accounts interacting with tweets for the first time even after a daily campaign extending five weeks.

d. Comments

Qualitatively, a significant proportion of replies to our study tweets were negative in nature. In response to the tweet that "smart investors do their homework and check their investment professional's background," for example, one user replied: "No. Smart investors run their own money!" In response to the tweet that "1 in 14 investment professionals have a record of misconduct," another user replied: "That doesn't include the ones NOT registered here. About 9/10 of those are fraudsters. You need to clamp down on those ones. All over social media." Many of these replies appear to have come from bot accounts which have subsequently been deleted since the study occurred.[11]

[11] Twitter conducted a systematic sweep of bot and fake accounts during the summer of 2018 around phase 2 of our study. See: https://www.washingtonpost.com/technology/2018/07/06/twitter-is-sweeping-out-fake-accounts-like-never-before-putting-user-growth-risk/?noredirect=on&utm_term=.6d64f861a642.

We assess whether a particular twitter account is spam-oriented by computing its "follower-to-following ratio", i.e., the number of users which follow the account divided by the number of users the account follows. An account with a lower ratio is considered to be more likely to be spam-oriented. Table 7 shows that the preponderance of replies came from spam-oriented accounts according to our measure. In particular, the preponderance of replies came from accounts with below the median follower-to-following ratio among accounts that interacted with that tweet (in terms of likes, replies, and retweets).[12]

However, negative replies did not appear to adversely impact the effect of tweets in terms of link clicks. Table 8 shows that the correlation between the number of replies and link clicks was positive across tweets in both phases. One reason may be because the twitter algorithm may have prioritized tweets based on number of replies without penalizing tweets for negative content in replies. Therefore, there appears to be some truth to the adage that "there's no such thing as bad publicity" within twitter campaigns.

III. Conclusion

In this study, we generate a number of findings relevant to academic research, regulation, financial education and outreach. First, there is some evidence that shorter, simpler calls to action worked best in promoting views and interactions with online messages. Second, there is also some evidence of greater (lower) engagement at the beginning (end) of the week as well as attrition in views and interactions over the five weeks of each study phase. However, this attrition was not with high turnover in accounts engaging tweets over the course of the study.

There are a number of avenues for further research. First, this study was not designed to test the effect of message length or complexity on behavior. Our findings suggest the need for further investigation into such effects controlling for message content, ability to see the full message on the viewer's screen, etc. In addition, this study represents the first step in a larger agenda to study the effects of messaging on encouraging beneficial financial behaviors such as reviewing disclosure documents, reporting fraud, etc. We hope to pursue further research employing bona-fide A/B testing to ensure true randomization and using promoted messages to reach a larger audience. Additional topics of interest include testing the effect of images in messaging to draw attention and promote financial behaviors.

[12] All accounts that replied to tweets 1 and 2 during phase 2 were subsequently deleted as of early August 2018 when we conducted this analysis.

Figure 1: Results over Time – Phase 1

The charts below show link clicks, engagements, and impressions per 1000 followers for weeks 1-5 of phase 1 of the study. The axis for impressions is on the left side, while the axis for engagements is on the right.

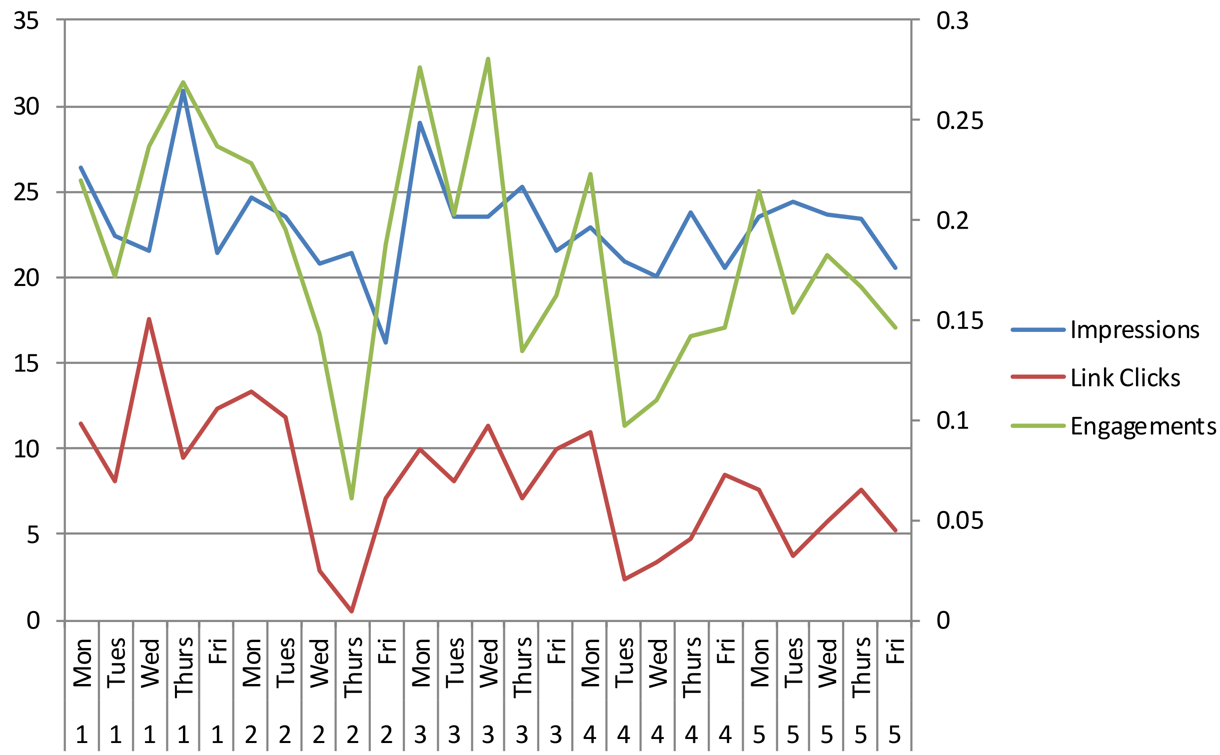

Figure 2: Results over Time – Phase 2

The charts below show link clicks, engagements, and impressions per 1000 followers for weeks 1-5 of phase 2 of the study. The axis for impressions is on the left side, while the axis for engagements is on the right.

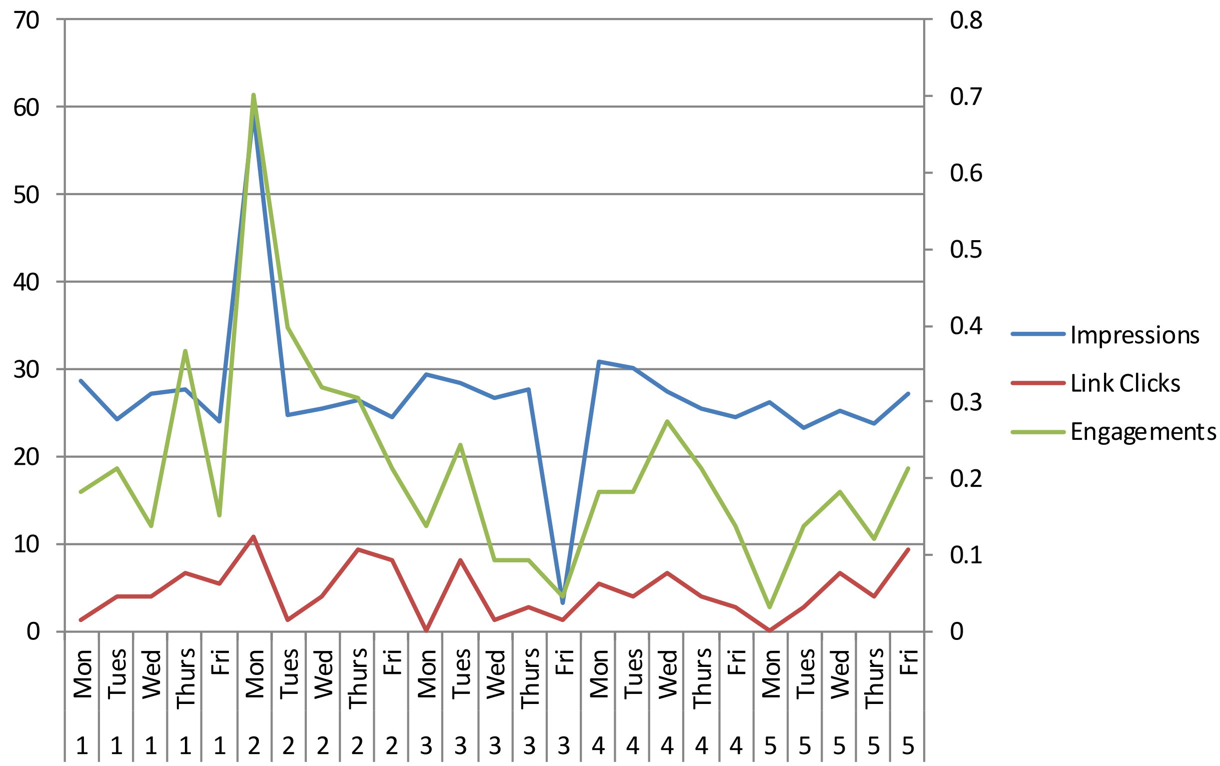

Table 1: Tweet Order

This table reports the order of tweets for phases 1 and 2 of the study. Week 3 of phase 2 was discarded because of a retweet error that Friday, while week 6 was added. Tweets during phase 1 were retweeted through SEC_News from Jan 22nd until Feb 23rd, 2018, while tweets during phase 2 were retweeted through SEC_Investor_Ed from April 2nd until May 4th plus June 11th-15th, 2018. The content of the tweets was as follows: 1.) Did you know? You can check the background of your investment professional here: 2.) Smart investors do their homework and check their investment professional's background. 3.) Studies suggest that 1 in 14 investment professionals have a record of misconduct! 4.) Fraud is often conducted by unregistered investment professionals! 5.) Avoid losing money to misconduct by an unlicensed investment professional!

Panel A: Phase 1

Week #	Mon	Tues	Wed	Thurs	Fri
1	2	3	1	4	5
2	4	1	2	5	3
3	3	4	5	2	1
4	1	5	4	3	2
5	5	2	3	1	4

Panel B: Phase 2

Week #	Mon	Tues	Wed	Thurs	Fri
1	3	2	4	5	1
2	4	5	2	1	3
~~3~~	~~1~~	~~4~~	~~5~~	~~3~~	~~2~~
4	2	3	1	4	5
5	5	1	3	2	4
6	1	4	5	3	2

Table 2: Tweet Averages Relative to Control

This table reports differences between average link clicks, engagements, and impressions per 1000 followers for tweets 2-5 relative to the control, tweet 1. P-values are reported below the differences of the averages. (***), (**), and (*) indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A: Phase 1

	Differences from Tweet 1		
Tweet	Links Clicks	Engagements	Impressions
2	-0.000208** 0.027	-0.000187 0.151	0.004506 0.677
3	-0.000191** 0.039	-0.000024 0.450	0.002057 0.580
4	-0.000159* 0.074	-0.000028 0.443	0.006883 0.763
5	-0.000204** 0.030	-0.000094 0.305	-0.001911 0.415

Panel B: Phase 2

	Differences from Tweet 1		
Tweet	Links Clicks	Engagements	Impressions
2	-0.000121 0.125	0.000000 0.500	-0.002246 0.435
3	-0.000046 0.332	0.000072 0.566	0.002180 0.572
4	0.000001 0.513	0.000501 0.883	0.039115 0.963
5	-0.000227** 0.012	0.000047 0.543	-0.001174 0.467

Table 3: Minimum and Maximum by Tweet

This table reports the minimum and maximum average link clicks, engagements, and impressions per 1000 followers across tweets 1-5 for both phases of the study. P-values for the minimum or maximum are reported below the average. (***), (**), and (*) indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A: Phase 1

Maximum

	Clicks	Engagements	Impressions
Tweet #	1	1	4
Value	0.0993*	0.197	23.9
P-val	0.0713	0.9365	0.8685

Minimum

	Clicks	Engagements	Impressions
Tweet #	2	2	5
Value	0.0578	0.159	22.2
P-val	0.8205	0.6543	0.8803

Panel B: Phase 2

Maximum

	Clicks	Engagements	Impressions
Tweet #	4	4	4
Value	0.0759	0.313	34.0
P-val	0.5127	0.9799	0.123

Minimum

	Clicks	Engagements	Impressions
Tweet #	5	2	2
Value	0.0304**	0.212	25.8
P-val	0.039	0.9799	0.7837

Table 4: Minimum and Maximum by Day-of-Week

This table reports the minimum and maximum average link clicks, engagements, and impressions per 1000 followers across days of the week for both phases of the study. P-values for the minimum or maximum are reported below the average. (***), (**), and (*) indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A: Phase 1

Maximum

	Clicks	Engagements	Impressions
Day	Monday	Monday	Monday
Value	0.0912	0.233*	25.3
P-val	0.273	0.0602	0.1634

Minimum

	Clicks	Engagements	Impressions
Day	Thursday	Thursday	Friday
Value	0.0504	0.155	20.0*
P-val	0.401	0.4811	0.0077

Panel B: Phase 2

Maximum

	Clicks	Engagements	Impressions
Day	Thursday	Thursday	Monday
Value	0.0757	0.276	35.1***
P-val	0.5515	0.9334	0.01

Minimum

	Clicks	Engagements	Impressions
Day	Tuesday	Friday	Friday
Value	0.0394	0.188	24.9
P-val	0.2508	0.6944	0.171

Table 5: Minimum and Maximum by Week

This table reports the minimum and maximum average link clicks, engagements, and impressions per 1000 followers across weeks 1-5 for both phases of the study. P-values for the minimum or maximum are reported below the average. (***), (**), and (*) indicate statistical significance at the 0.01, 0.05 and 0.10 levels, respectively.

Panel A: Phase 1

Maximum

	Clicks	Engagements	Impressions
Week #	1	1	3
Value	0.101**	0.227	24.6
P-val	0.04	0.1286	0.5038

Minimum

	Clicks	Engagements	Impressions
Week #	5	4	2
Value	0.0511	0.144	21.3
P-val	0.4208	0.1925	0.3325

Panel B: Phase 2

Maximum

	Clicks	Engagements	Impressions
Week #	2	2	2
Value	0.0763	0.387**	32.1
P-val	0.4574	0.0138	0.8449

Minimum

	Clicks	Engagements	Impressions
Week #	1	5	5
Value	0.0489	0.137**	25.1
P-val	0.8531	0.0228	0.3006

Table 6: Repeat Interactions

This table reports the average number of interactions (i.e., replies, retweets, and likes) with our study tweets accounts accounts which interacted with the tweets from our study.

	Total Users	Mean Interactions	Median Interactions	95th Percentile Interactions	% of Users with One Interaction	% of Users Interacting in Both Phases
Phase 1	137	1.47	1	3.2	74%	
Phase 2	69	1.45	1	3.6	70%	
Combined	202	1.50	1	3.9	70%	2.5%

Table 7: Follower-to-Following Ratio for Replies

This table reports the number of accounts with below median "follower-to-following ratio" (i.e., the number of twitter users which follow the account divided by the number of users the accounts follows) by tweet and phase for accounts which replied to our study tweets. The median is computed among accounts that interacted with that tweet (in terms of likes, replies, and retweets). Any accounts which replied to tweets 1 and 2 during phase 2 were subsequently deleted as of early August 2018 when we conducted this analysis.

	% of Comments by Users with Below-Median Ratio	
Tweet Number	Phase 1	Phase 2
1	100%	-
2	57%	-
3	67%	100%
4	88%	100%
5	100%	75%
Average	**76%**	**83%**

Table 8: Link Click Correlations

This table reports the correlation coefficient between link clicks and a number of dependent variables (impressions, detail expands, replies, likes, and retweets) across tweets in our study.

	Impressions	Detail Expands	Replies	Likes	Retweets
Phase 1	25%	12%	2%	15%	43%
Phase 2	42%	18%	31%	49%	12%

References

Bartov, Eli, Lucile Faurel, and Partha Mohanram, 2018, "Can Twitter Help Predict Firm-Level Earnings and Stock Returns?" *Accounting Review*, 93, 25-57.

Bauer, Rob, Inka Eberhardt, and Paul Smeets, 2017, "Financial Incentives Beat Social Norms: A Field Experiment on Retirement Information Search," *Working Paper, available at: https://ssrn.com/abstract=3023943 or http://dx.doi.org/10.2139/ssrn.3023943*

Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian, Katherine L. Milkman, 2015, "The Effect of Providing Peer Information on Retirement Savings Decisions" *Journal of Finance,* 70(3), 1161-1201.

Bhargava, Saurabh and Dayanand Manoli, 2014, "Why Are Benefits Left on the Table? Assessing the Role of Information, Complexity, and Stigma on Take-Up with an IRS Field Experiment," *Working Paper, Center for Behavioral Decision Making Research, Carnegie Mellon University*.

Bollen, Johan, Huina Mao, and Xiao-Jun Zeng, 2010, "Twitter Mood Predicts the Stock Market," *Journal of Computational Science*, 2, 1-8.

Burke, Jeremy, and Angela A. Hung, 2015, "Trust and Financial Advice," *RAND Working Paper WR-1075*.

Cialdini, Robert B. and Noah J. Goldstein, 2004, "Social Influence: Compliance and Conformity*" Annual Review of Psychology,* 55(1), 591-621.

Chong, Alberto, Dean Karlan, Jeremy Shapiro, and Jonathan Zinman, 2015, "(Ineffective) Messages to Encourage Recycling: Evidence from a Randomized Evaluation in Peru", *World Bank Economic Review,* 29, 180-206.

Dimofte, Claudiu V., Ronald C. Goodstein, and Anne M. Brumbaugh, 2014, "A Social Identity Perspective on Aspirational Advertising: Implicit Threats to Collective Self-Esteem and Strategies to Overcome Them," *Journal of Consumer Psychology,* 25(3), 416-430.

Gerber, Alan S., Donald P. Green and Christopher W. Larimer, 2008, "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment," *American Political Science Review*, 102(1), 33-48.

Gu, Chen, and Alexander Kurov, 2018, "Informational Role of Social Media: Evidence from Twitter Sentiment," *Working paper.*

Gu, Chen and Denghui Chen, 2018, "Does Investor Attention Foretell Stock Trading Activities? Evidence from Twitter Attention," *Working paper, available at: https://ssrn.com/abstract=3219221 or http://dx.doi.org/10.2139/ssrn.3219221*

Hallsworth, Michael, John A. List, Robert D. Metcalfe, Ivo Vlaev, 2017, "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance," *Journal of Public Economics,* 148, 14-31.

Hwang, Byoung-Hyoun and Hugh Hoikwang Kim, 2017, "It Pays to Write Well," *Journal of Financial Economics*, 124(2), 373-394.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman, 2016, "Getting to the Top of Mind: How Reminders Increase Saving" *Management Science*, 62(12), 3393-3672

Lawrence, Alastair, 2013, "Individual Investors and Financial Disclosure," *Journal of Accounting and Economics,* 56(1), 130-147.

Lehavy, Reuven, Feng Li, Kenneth Merkley, 2011, "The Effect of Annual Report Readability on Analyst Following and The Properties of Their Earnings Forecasts," *Accounting Review,* 86(3), 1087–1115.

Loughran, Tim, and Bill McDonald, 2014, "Measuring Readability in Financial Disclosures," *Journal of Finance,* 69(4), 1643–1671.

Madlener, Reinhard, and Blake Alcott, 2009, "Energy Rebound and Economic Growth: A Review of the Main Issues and Research Needs," *Energy,* 34(3), 370-376.

Mao, Huina, Scott Counts, and Johan Bollen, 2015, "Quantifying the Effects of Online Bullishness on International Financial Markets," *ECB Working Paper.*

Mao Yuexin, Bing Wang, Wei Wei, Benyuan Liu, 2012, "Correlating S&P 500 Stocks with Twitter Data" *Working Paper.*

Oliveira, Nuno, Paulo Cortez, and Nelson Areal, 2017, "The Impact of Microblogging Data for Stock Market Prediction: Using Twitter to Predict Returns, Volatility, Trading Volume and Survey Sentiment Indices," *Expert Systems with Applications*, 73, 125-144.

Schultz, P. Wesley, 1999, "Changing Behavior with Normative Feedback Interventions: A Field Experiment on Curbside Recycling," *Basic and Applied Social Psychology*, 21, 25-36.

Tversky, Amos, and Daniel Kahneman, 1992, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297-323.

Verhallen, Pieter, Elisabeth Brüggen, Thomas Post and Gaby Odekerken-Schröder, 2018, "Norms in Behavioral Interventions: Peer or Anchoring Effects?" Working paper, a*vailable at: https://ssrn.com/abstract=3098028 or http://dx.doi.org/10.2139/ssrn.3098028*

U.S. Social and Behavioral Sciences Team, 2015, "SBST Annual Report."

You, Haifeng, and Xiao-jun Zhang, 2009, "Financial Reporting Complexity and Investor Underreaction to 10-K Information," *Review of Accounting Studies,* 14(4), 559–586.