Short and Distort

Joshua Mitts*

August 17, 2018

Abstract

Pseudonymous attacks on public companies are followed by stock price declines and sharp reversals. I find these patterns are likely driven by manipulative stock options trading by pseudonymous authors. Among 1,720 pseudonymous attacks on mid- and large-cap firms from 2010-2017, I identify over $20.1 billion of mispricing. Reputation theory suggests these reversals persist because pseudonymity allows manipulators to switch identities without accountability. Using stylometric analysis, I show that pseudonymous authors exploit the perception that they are trustworthy, only to switch identities after losing credibility with the market.

*Associate Professor of Law, Columbia Law School. I am grateful to Ian Ayres, Lucian Bebchuk, Bernie Black, Alon Brav, Anthony Casey, Emiliano Catan, Albert Choi, John C. Coffee, Jr., Martijn Cremers, Ofer Eldar, Jill Fisch, Elisabeth de Fontenay, Slava Fos, Ed Fox, Merritt Fox, Jens Frankenreiter, Jesse Fried, Jeffrey Gordon, Arpit Gupta, Ernan Haruvy, Scott Hirst, Colleen Honigsberg, Gur Huberman, Robert J. Jackson, Jr., Wei Jiang, Donald Langevoort, Jonathan Macey, Justin McCrary, Tamar Mitts, Ed Morrison, Darius Palia, Jim Park, David Pozen, Alex Raskolnikov, Gabriel Rauterberg, Adriana Robertson, Roberta Romano, Natasha Sarin, Michael Simkovic, Eric Talley, and Robert Weber for very insightful comments.

1 Introduction

Pseudonymity—writing under a fictitious identity—encourages expression of controversial ideas but also reduces accountability (Publius, 1787). While anonymous political speech enjoys strong protections under the U.S. Constitution,[1] in financial markets pseudonymity facilitates profitable manipulation of stock prices. I show that pseudonymous authors publish negative rumors about public companies that lead to significant short-term trading profits—and sharp reversals of the stock price decline. When markets realize that the pseudonymous author is spreading baseless rumors, the author switches to a new pseudonym, repeating the pattern.

In this article, I conduct the first systematic study of pseudonymous "short-and-distort" attacks on public companies. I examine 2,900 attacks on mid- and large-cap firms published on a website, Seeking Alpha, and show that pseudonymous ones are accompanied by a decline in the target's stock price followed by a sharp reversal, leading to over $20.1 billion in mispricing. To address selection concerns, I employ propensity-score matching between pseudonymous and real-name attacks. I then use a triple-difference design to show contemporaneous stock options trading with publication. While I cannot identify who exactly is trading, the universe of potential traders is small: only the pseudonymous author, his/her tippees, and possibly the Seeking Alpha editorial staff know that an attack article is forthcoming.[2] A textual analysis shows that provocative article content is unlikely to explain these price reversals. Finally, I show that market makers widen bid/ask spreads in response.

Why do these patterns persist in highly liquid, otherwise efficient markets? Drawing on Benabou & Laroque (1992), I show that pseudonymity undermines reputational mechanisms that discredit systematic manipulators. Like Publius (1787), a fictitious identity can shield the dissemination of controversial truth as well as falsehood, causing market participants to listen to pseudonymous speech. I show that pseudonymous authors exploit the perception that they are trustworthy, only to switch identities after losing credibility with the market. Unbridled

[1] In the words of Justice Stevens, "Anonymity is a shield from the tyranny of the majority." McIntyre v. Ohio Elections Com'n, 514 U.S. 334, 357 (1995) (citing J. S. Mill, On Liberty, *in* On Liberty and Considerations on Representative Government 1, 3-4 (R. McCallum ed. 1947)).

[2] Seeking Alpha strictly prohibits editors from trading ahead of a forthcoming article, see https://seekingalpha.com/page/seeking-alpha-conduct-and-investment-policy, and there is no evidence that they are doing so.

candor is a double-edged sword, allowing authors to mislead without reputational penalties. Pseudonymity thus implicates a difficult cost-benefit tradeoff, encouraging open dialogue but making it harder to hold fictitious identities accountable for manipulative speech.

Pseudonymous attacks pose unique challenges for securities law. A factual misstatement made with scienter in connection with the purchase or sale of a security gives rise to liability under Section 10(b) of the Securities Exchange Act of 1934. But pseudonymous attacks often consist of murky opinions rather than express factual claims, and it is difficult to establish intent to deceive without knowing the identity of the pseudonymous author. Moreover, it is difficult to establish that open-market transactions constitute market manipulation in violation of Section 9 of the Securities Act of 1933, even with proof of intent to distort prices.

The key to this story is online intermediaries like Seeking Alpha, who secure fictitious accounts and keep them from being hijacked by anonymous impersonators. As gatekeepers of the link between pseudonymous accounts and underlying authors, these intermediaries are well-suited to punish systematic manipulators without chilling pseudonymity itself. The Securities and Exchange Commission should hold these intermediaries accountable for tolerating market manipulation via pseudonymous attacks.

2 Anecdotal Example and Literature

I begin with an anecdotal example of a pseudonymous attack. Insulet Corporation (NASDAQ: PODD) is a publicly traded medical device manufacturer based in Billerica, Massachusetts with a market value of $5.8 billion as of May 2018. Insulet manufacturers the Omnipod insulin pump, which gives diabetics an alternative to multiple daily insulin injections. On November 29, 2016, an article about Insulet was published on the website Seeking Alpha, a platform for investors to author article-length blog posts about public companies.[3] The article had a salacious title: *Insulet Investors Being Kept In The Dark, CEO Alleged To Encourage Questionable Sales Techniques*, and claimed (1) to have "obtained evidence of yet another whistleblower payoff," (2) the "CEO allegedly directed employees to bribe physicians" and (3)

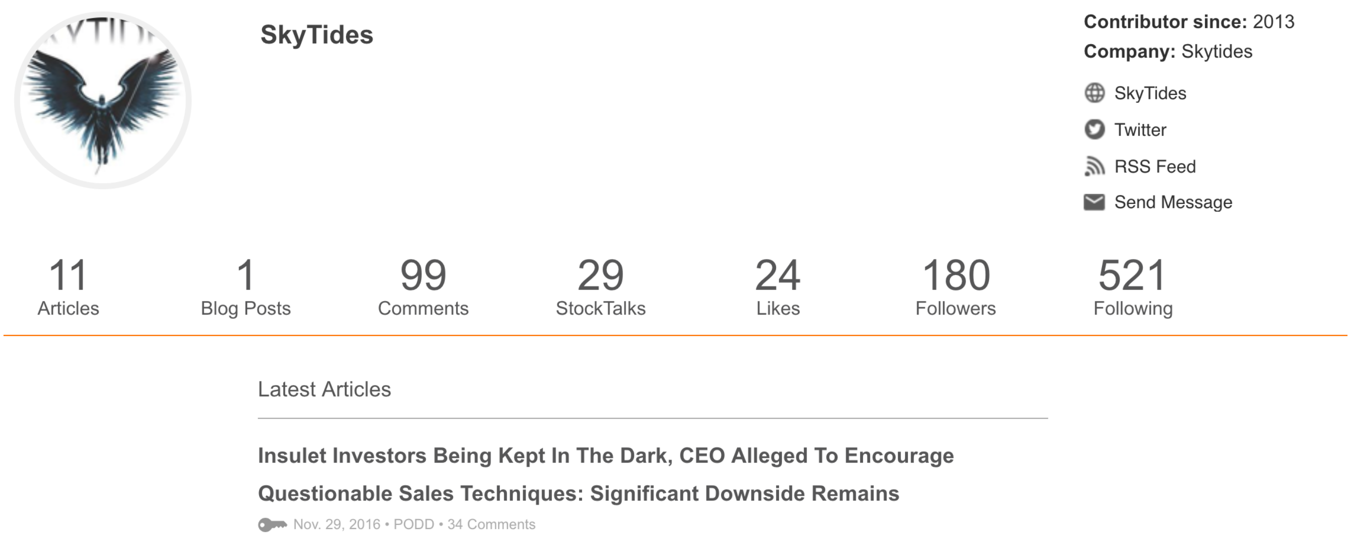
[3] http://www.seekingalpha.com

"multiple sell-side analysts claimed CEO deceived investors by not fully disclosing the extent of Omnipod product defects and prior management's fraudulent acts."[4]

The article was written by an author named "SkyTides," a pen name for a pseudonymous blogger on Seeking Alpha. The platform proudly encourages pseudonymity, pointing out that "regulations at their workplace or other factors" prevent "some contributors [from] revealing their real names. In addition, many well-known, veteran stock market bloggers (some of the finest, in fact) write under a pseudonym."[5] The profile page for SkyTides reveals nothing about who this author actually is, as shown in Figure 1.

Figure 1: Seeking Alpha Profile of SkyTides

This image is the Seeking Alpha profile page of SkyTides, a pseudonymous author who attacked Insulet.



One might assume that markets would pay little attention to a pseudonymous author like SkyTides. After all, unlike an identifiable author posting under a real name, it is hard to hold SkyTides accountable for authoring misleading or inaccurate information. These kinds of pseudonymous postings seem like a quintessential example of "cheap talk" lacking credibility (Farrell & Rabin, 1996): pseudonymity makes it virtually costless for SkyTides to lie, so rational investors should ascribe little, if any, weight to what SkyTides says.

However, immediately following the posting of SkyTides' article on November 29, 2016,

[4] https://seekingalpha.com/article/4026931-insulet-investors-kept-dark-ceo-alleged-encourage-questionable
[5] https://seekingalpha.com/page/policy_pseudonymous_contributors

the price of Insulet's stock fell by over 7% from $35.21 on November 28 (the day before the article's publication) to $32.77 on December 1 (two days after the article's publication). One might conclude that SkyTides was simply right – perhaps Insulet had some serious problems, and the market recognized this by bidding down the price of Insulet's stock. Indeed, SkyTides proudly touted this decline on their homepage.[6] But then a curious thing happened: Insulet's price climbed right back up on December 5th (four trading days after publication), and rose higher than where it closed before SkyTides's article was posted. Figure 2 shows how Insulet's stock price displays a "V" pattern centered on the publication of SkyTides' article on 11/29:

Figure 2: Insulet Corporation (NASDAQ: PODD) Stock Price

This image shows the stock price of Insulet Corporation from November 21, 2016 to December 14, 2016. The stock price graph is taken from TradingView, www.tradingview.com.

A decline of over 7% is highly unlikely to have been caused by random chance. And there is evidence that Insulet's stock price was subjected to manipulative options trading alongside the publication of the article. Put and call options are contracts that allow investors to make bets that a company's stock price will rise or fall, respectively. These bets are high risk, high

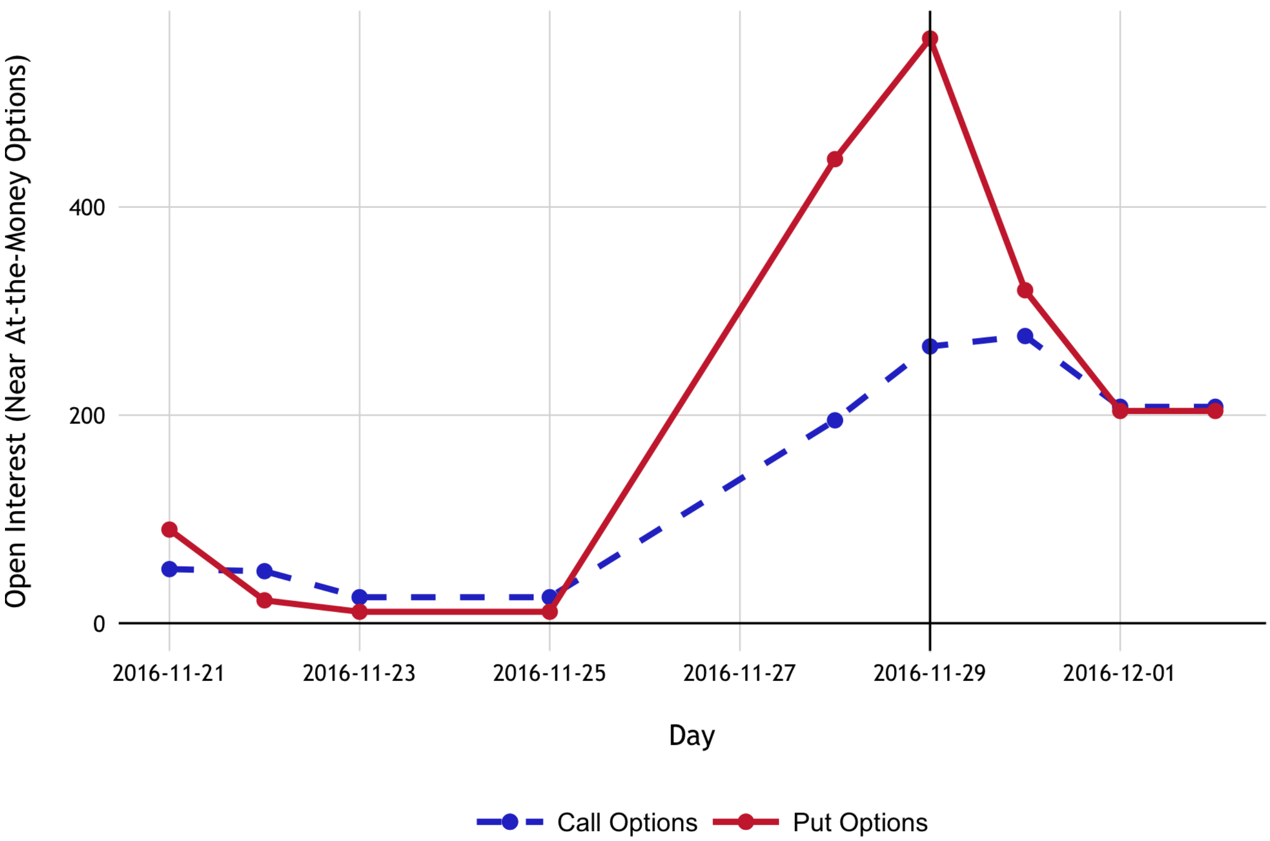
[6]http://www.skytides.com/

reward: if the stock price goes up, the value of a call option increases by a lot, but if it goes down, the call option becomes virtually worthless. Trading in options suggests someone has information about which way the stock price will go, and has served as a basis for insider trading cases brought by the SEC (Chakravarty *et al.* , 2004; Meulbroek, 1992).

Figure 3 plots the number of outstanding put and call option contracts[7] on Insulet's stock in the days before and after the publication of SkyTides' article: Figure 3 shows that there was

Figure 3: Insulet Corporation (NASDAQ: PODD) Near-ATM Options: 11/21/16-12/2/16

This image shows the open interest of call and put options written on Insulet Corporation. These options are nearly at-the-money, i.e., they have an absolute delta between 0.45 and 0.55.



a large purchase of put options the day before the article's publication[8], which pay off if the stock price declines (which it did), and sold those put options immediately thereafter, which

[7]These options are nearly "at-the-money," i.e., they have an absolute delta between 0.45 and 0.55.
[8]Open interest is lagging by one day, so options reported on 11/29 were actually purchased on 11/28.

would cause the stock price to rise (which it did). That kind of well-timed options trading suggests that someone knew the article was about to be published, and that the price would revert to its prior level thereafter because the article did not contain sufficient information to bring about a downward revision in the price of the magnitude observed on the day of publication. Indeed, while the subsequent rise in the price could have been driven by public arbitrageurs, nobody knew the article was forthcoming except the author, his or her tippees, or possibly the Seeking Alpha editorial staff,[9] so the put options are especially suspicious.

SkyTides-Insulet is hardly an isolated case. Short sellers have increasingly embraced this kind of pseudonymous online activism. Two lawyers at Ropes & Gray LLP recently wrote that "pseudonymous online hit pieces against public companies have become an increasingly common and effective form of short activism,"[10] and pointed to several substantial price declines in the wake of pseudonymous attacks. And three lawyers at DLA Piper recently discussed the cases of Chromadex Inc., which was attacked by a pseudonymous short seller and lost $100 million of market capitalization in a single day, and Cemtrex, which attempted to sue the pseudonym "Richard Pearson" but failed for lack of service of process.[11]

Are these short attacks informative or manipulative? The SEC adopted Regulation SHO out of a concern that some kind of shorting could be "manipulative or abusive."[12] A large literature in finance contends that these legal constraints reduce price efficiency (Beber & Pagano, 2013; Boehmer *et al.* , 2013; Comerton-Forde & Putniņš, 2008; Saffi & Sigurdsson, 2010). Zhao (2018) identifies a correlation between being targeted by activist short sellers and firm characteristics like overvaluation and uncertainty. Wong & Zhao (2017) find that the targets of short activism experience a subsequent decline in investment, financing and payouts. Campbell *et al.* (2017) find that between position disclosures by activist short sellers are linked to differences in short-run returns and are not interpreted by investors as evidence of bias.[13]

---

[9]For more on Seeking Alpha's policy against front-running by editors, see note 2 *supra*.

[10]https://corpgov.law.harvard.edu/2017/11/27/short-activism-the-rise-in-pseudonymous-online-short-attacks/

[11]https://www.dlapiper.com/~/media/files/people/weiner-perrie/weinerweberhsu.pdf

[12]https://www.sec.gov/divisions/marketreg/mrfaqregsho1204.htm

[13]In my data, I find that position disclosure is highly correlated with firm characteristics already included in the propensity-score matching, including idiosyncratic volatility, Amihud (2002) illiquidity, market capitalization, total assets and liabilities, and industry group. The informed trading literature suggests that the ability to accumulate a position without moving the market price in the direction of one's information turns on market impact, i.e., position

In a systematic study of market manipulation, Fox *et al.* (2018) write that "the core harm of a manipulation will actually depend on the speed and nature of [a] price 'correction'." Following this definition, I show that pseudonymous attacks are accompanied by a decline in the price of target firms followed by a sharp reversal. In addition, I use a triple-differences design to show that pseudonymous articles are accompanied by well-timed options trading. On the day of publication, the open interest and volume of put options written on the target of a pseudonymous article are higher than call options. While I cannot prove that the pseudonymous author is trading, the universe of potential traders is small: only that author, his or her tippees, or possibly the Seeking Alpha editorial staff,[14] know an article is forthcoming. Because open interest is lagged by one day, these options were purchased prior to publication.

During the second to fifth day following publication, the open interest and volume of call options written on the target of a pseudonymous article are higher than put options. Both calls and puts follow parallel trends in the preceding days, strengthening the causal interpretation of the divergence in open interest and volume. In addition, following Cremers & Weinbaum (2010), I show that these periods are characterized by deviations from put-call parity that are indicative of informed trading. Moreover, a textual analysis suggests that provocative article content is unlikely to be driving these price reversals. The words and phrases correlated with pseudonymous authorship do not refer to fraud or similar evocative improprieties.

This empirical pattern is consistent with the characterization of "misstatement manipulation" in Fox *et al.* (2018) as a form of informed trading (pp. 112-13):

> Misstatement manipulation is actually a special kind of informed trading. Just like a corporate insider who trades on confidential information that he obtained from his employer, the misstatement manipulator makes his purchases on the basis of something that he knows and the market does not: the falsity of the price-depressing misstatement for which he is responsible.

Fox *et al.* (2018) point out that "liquidity suppliers will increase their spreads to compensate for the prospect of losing money to misstatement manipulators . . . the decrease in liquidity may be the more serious socially negative effect."

disclosure is likely to proxy for illiquidity.

[14]For more on Seeking Alpha's policy against front-running by editors, see note 2 *supra*.

I directly test this proposition by examining how market makers adjust bid/ask spreads in anticipation of informed buying during the reversal period. While the publication of the article comes as a surprise to market makers (so they cannot adjust spreads to compensate for the losses that will be incurred by buying prior to the impending price decline), they can anticipate the possibility of selling to an "informed buyer" who purchases in anticipation of an impending post-publication price correction. I show that a stronger negative cumulative abnormal return on the publication day is linked to an increase in bid/ask spreads until two days post-publication, when call options trading is expected to commence.

This article contributes to an emerging literature in finance on the link between websites like Seeking Alpha and the financial markets. Kogan *et al.* (2018) show that the publication of "fake news" on social media, blogs and similar outlets is followed by temporary price impact and reversals for small firms, but not for mid- or large-cap firms. The analysis in Kogan *et al.* (2018) complements this project by examining the publication of factually false articles rather than opinion pieces by short sellers. And while they do not distinguish between pseudonymous vs. non-pseudonymous authors nor consider options trading, Kogan *et al.* (2018) show that managers of small- and mid-cap firms may be engaging in market manipulation by issuing press releases, filing Form 8-K disclosures and engaging in insider trading.

My findings also speak to the large literature on informed trading in options markets (Chakravarty *et al.* , 2004; Easley *et al.* , 1998). Cremers & Weinbaum (2010) find that deviations from put-call parity predict future stock returns: stocks with more expensive calls outperform stocks with more expensive puts. An *et al.* (2014) examine the joint cross section of stocks and options, and find that implied volatility predicts future stock returns, as suggested by a rational model of informed trading. Consistent with this literature, I find that the targets of pseudonymous short attacks undergo similar deviations from put-call parity in the days accompanying the attack, suggesting the presence of informed trading.

Theoretical work has long considered how manipulators may depress prices by releasing negative information, profiting on the decline and subsequent price reversal. Vila (1989) articulates this basic strategy in a game-theoretic framework. Van Bommel (2003) considers how imprecise rumors can lead to systematic price declines with occasional overshooting, allowing even honest rumormongers to profit from price reversals. Van Bommel (2003) relies on stylized assumptions, including that rumors are shared with a small group of "followers" subject to a fixed trading capacity, and only slowly diffused to public arbitrageurs. These assumptions are violated by the publication of pseudonymous attacks on a public website like Seeking Alpha.

In a survey of information-based manipulation, Putniņš (2012) points out that reputation can limit manipulation to one shot: "if market participants are able to deduce that false information originated from a manipulator, the manipulator will quickly be discredited and the manipulation strategy will cease to be profitable." Benabou & Laroque (1992) show that when private signals are noisy, market participants cannot infer with certainty that a price reversal is indicative of a manipulator, as the author might simply have relied on incorrect facts. But few pseudonymous attacks seem to arise out of honest mistakes.

Another explanation for the persistence of price reversals in a repeated setting, consistent with Benabou & Laroque (1992), is that pseudonymity allows an author to switch fictitious identities, preventing market participants from discrediting systematic manipulators. Moreover, unlike spam stock tips or message boards of the late 1990s, pen names today are often employed to disseminate truth as well as falsity, both because of employer social media policies as well as to protect authors from frivolous litigation and legal uncertainty. Renaulta (2018) examine over 7 million posts on Twitter and find that a burst of social media activity about small-cap companies is followed by a price increase and subsequent reversal over the next week. Consistent with my findings, Renaulta (2018) identifies the pseudonymity and possible truthfulness of Twitter content as facilitating this sort of fraud.

This theoretical framework yields several testable predictions. First, pseudonymous authors should engage in informed trading when they are perceived as trustworthy by the market, such as when they have no history or the author has had few reversals in the past. Second, pseudonymous authors should "disappear" after the market realizes fraud is taking place, so that they can switch to a new identity. Finally, switching pseudonymous identities should leave subtle traces of writing style which can be detected using linguistic stylometry. The following Section 3 presents evidence consistent with all of these predictions.

3 Empirical Study

3.1 Data and Sample Construction

I begin by collecting all articles published on Seeking Alpha under the category "Short Ideas" from January 1, 2010 to December 31, 2017. That category contains all articles which advocate taking a short position in one or more firms. Seeking Alpha provides the exact date and time that the article was published, as well as the ticker of the firm(s) that are the subject of the article. This yields an initial sample of 14,730 articles.

To determine which authors are pseudonymous, I hired workers from the crowdsourcing website Figure Eight. I asked workers to look up the name of the author on Seeking Alpha, determine whether he or she is pseudonymous based on the absence of personally identifiable biographical information in their Seeking Alpha profile. For each author, I had three workers evaluate his or her profile, and I coded an author as pseudonymous if and only if all three authors agreed that the profile did not refer to an identifiable individual. In addition, I manually verified and corrected a few sporadic errors in the coding. Table 1 shows ten example authors from the pseudonymous and non-pseudonymous groups, respectively.

[Table 1]

To accurately measure trading behavior around the publication of attacks, I remove any article published about the same firm within 7 calendar days of a prior article. There are a few firms (like Tesla) which are the subject of near-daily attacks by short sellers. In that case, it is difficult to view the publication of each additional article as a new informative attack rather than a reiteration of what is already known. Moreover, it is important to verify that the results are not driven by these arguably pathological cases of incessant publications about the same firm rather than publications which bring new information to the market. This yields 9,121 articles about 2,311 publicly traded firms.

In addition, because this study depends heavily on market participants rapidly responding to and trading on the basis of information publicly disclosed in these articles, I limit my primary analysis to mid-cap and large-cap firms with at least $2 billion in market capitalization. The inclusion of small- and micro-cap stocks adds noise to the data, as prices are often much slower to respond, and their relative illiquidity and lower nominal prices leads to much greater nominal volatility of returns. This yields 4,785 articles about 837 publicly traded firms.

For each of these firm-article pairs, I obtain standard characteristics from Compustat like market value of equity, total assets, total liabilities and net income for the year preceding the article, and derive the Amihud (2002) illiquidity measure and idiosyncratic volatility using daily returns over the period $[t_0 - 120, t_0 - 7]$, where t_0 is the date of article publication. Summary statistics on my primary dataset are presented in Table 2.

[Table 2]

Which firms are targeted by pseudonymous authors? Table 3 considers predictors of pseudonymous authorship among the entire sample of 4,785 firm-articles.

[Table 3]

As Table 3 shows, pseudonymous targets tend to be slightly smaller and less profitable than real-name targets, but indistinguishable in terms of assets and liabilities. There are sector-specific differences, e.g., consumer durables and apparel are more likely to be targeted by pseudonymous authors, whereas retailing and software & services are more likely to be targeted by real-name authors. The following Section details the use of propensity score matching to obtain a sample that is balanced on these observable characteristics.

3.2 Propensity-Score Matching

A naive comparison of market reactions to pseudonymous to non-pseudonymous articles is subject to the critique that these reactions may be driven by unobserved differences between firms which are the targets of these articles. To be sure, this concern is less compelling in this kind of event-study setting involving high-frequency outcomes like price changes in the days following the publication of an blog post attacking a publicly traded company. To further mitigate selection concerns, I employ a matched design to ensure that I compare firms which are similar as possible on observable characteristics. I match pseudonymous and non-pseudonymous articles on the following firm and article characteristics: (1) market value of equity; (2) total assets; (3) total liabilities; (4) net income; (5) Amihud (2002) illiquidity; (6) the idiosyncratic volatility of the firm's stock; (7) GICS Industry Group code; and (8) the publication hour of the article, which adjusts for time-varying market liquidity conditions.

I present my results using nearest-neighbor matching, which yields a weighted sample of 2,900 article-firms. The results of a balance test on these covariates are given in Table 4.

[Table 4]

Table 4 shows that the treatment and control groups are balanced across all of these characteristics. A t-test of each of the variables yields p-values that are all above 5%, indicating the differences in means are not statistically significant As additional evidence that the two samples are balanced on these characteristics, Figure 7 in the Online Appendix presents the density of the propensity score between the treatment and control groups for the single-neighbor matching. As Figure 7 shows, the two groups have very similar densities.

3.3 Abnormal Returns to Article Publication

I begin my analysis by comparing cumulative abnormal returns to between pseudonymous and non-pseudonymous attacks on public companies. I fit a standard four-factor model of expected returns by estimating the following regression for each of the articles in my dataset by ordinary least squares on daily returns over the interval $[t_0 - 120, t_0 - 7)$ in calendar days (approximately $[t_0 - 85, t_0 - 5]$ in trading days), where t_0 is the date of article publication:

$$r_{i,t} - r_{f,t} = \beta_{i,0} + \beta_{i,1} m_t + \beta_{i,2} smb_t + \beta_{i,3} hml_t + \beta_{i,4} umd_t + \epsilon_{i,t}$$

where $r_{i,t}$ is the log return on the common stock of firm i on day t, $r_{f,t}$ is the log risk-free rate on day t, m_t is the log return on the market on day t, smb_t is the log return on the Fama-French small-minus-big portfolio on day t, hml_t is the log return on the Fama-French high-minus-low portfolio on day t, umd_t is the log return on the winners-minus-losers momentum portfolio (Carhart, 1991) on day t, and $\epsilon_{i,t}$ is a random error term.

Next, I obtain daily abnormal log returns by subtracting the predicted values given by this model from the actual returns for each day in the interval $[t_0 - 5, t_0 + 5]$ in trading days, where, as before, t_0 is the date of article publication:

$$\alpha_{i,t} = r_{i,t} - r_{f,t} - (\beta_{i,0} + \beta_{i,1} m_t + \beta_{i,2} smb_t + \beta_{i,3} hml_t + \beta_{i,4} umd_t)$$

Finally, I derive the cumulative abnormal log return from day t to day τ for firm-article i written by author j by summing the daily log abnormal returns:

$$car_{i,j,t,\tau} = \sum_{k=t}^{\tau} \alpha_{i,k}$$

Figure 4 plots $car_{i,j,t_0-5,\tau}$ for pseudonymous and non-pseudonymous articles with $\tau \in (t-5, t+5]$ in trading days. As Figure 4 shows, both pseudonymous and non-pseudonymous articles are accompanied by negative cumulative abnormal returns on the order of .01 log points, i.e., approximately 1 percentage point. There is little difference in the cumulative abnormal log return between pseudonymous and non-pseudonymous articles in the $[t_0 - 4, t_0 - 1]$ window: both groups experience a parallel decline prior to publication of the article, which is likely driven by general negative sentiment in the market.[15]

However, pseudonymous articles decline further on the day of publication (t_0) and display a sharp pattern of reversal over the $[t_0 + 2, t_0 + 5]$ window, with returns increasing from -0.0106 to -0.0073 from $t_0 + 2$ to $t_0 + 5$, a difference of 0.33 log points or approximately 31.1% in relative terms, from day 1 to day 5 following publication.

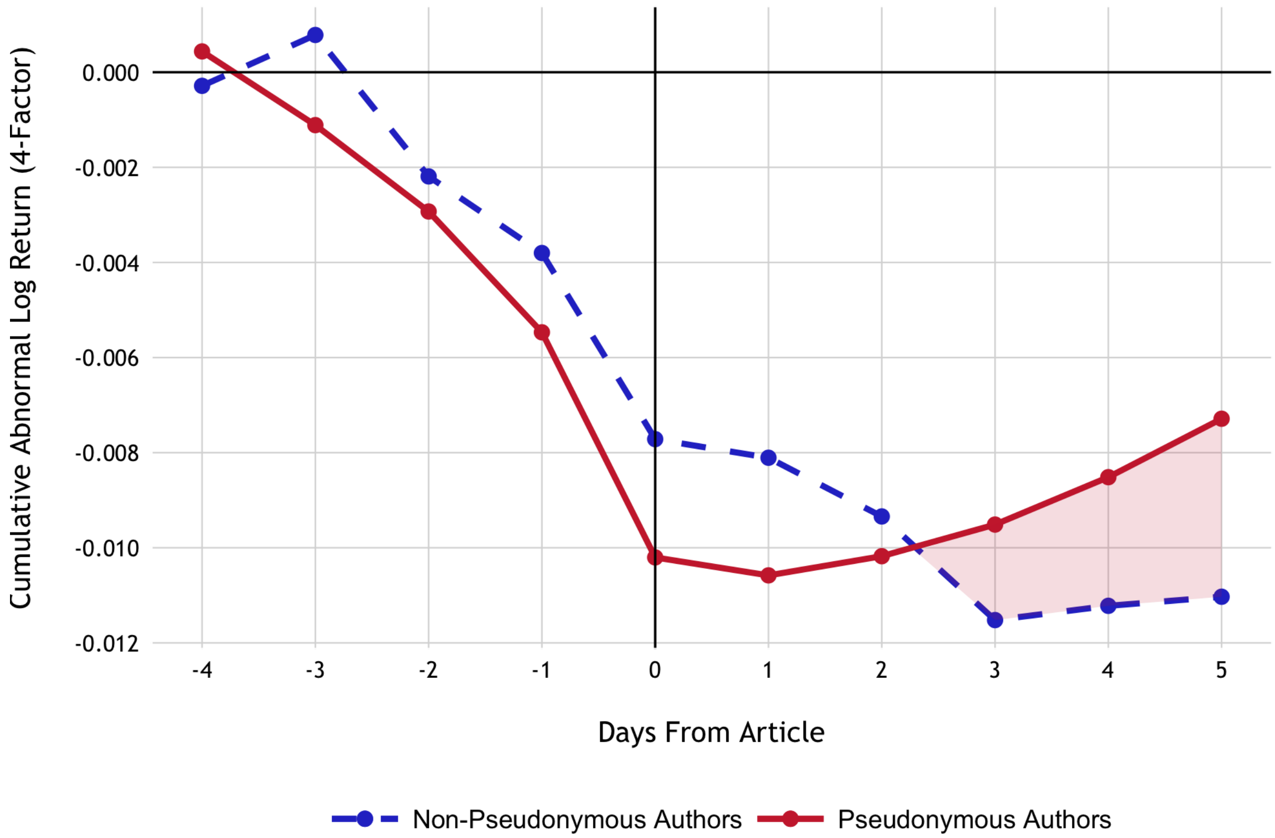
3.4 Stock Price Reversals

I systematically test whether pseudonymous articles are followed by greater stock-price reversals by estimating several different regression models on my data. I begin by implementing an overreaction measure following Tetlock (2011), regressing the cumulative abnormal return over the $[t_0 + 2, t_0 + 5]$ interval on the cumulative abnormal return over the $[t_0 - 1, t_0 + 1]$ interval. Unlike the average differences in Figure 4, a significant negative coefficient indicates

[15] Indeed, short attacks are often written in response to prior negative news or other negative sentiment, which is why it is difficult to identify the causal effects of these attacks (Zhao, 2018). My design compares pseudonymous to non-pseudonymous articles, where, as Figure 4 shows, the two groups follow parallel, albeit declining, pre-trends.

Figure 4: Daily Cumulative Abnormal Returns to Pseudonymous Attacks

This figure shows the average daily cumulative abnormal log returns from a four-factor Fama-French model to the publication of an attack article over the window $(t_0 - 5, t_0 + 5]$, where t_0 is the publication date of the article, for pseudonymous and non-pseudonymous authors separately. As the figure shows, articles published by pseudonymous articles are followed by a price reversal that is indicated by the shaded region.



a pair-wise negative correlation between the abnormal return in these two periods: the further prices fall after publication of the article, the higher they rise afterward.

The key question is whether price reversals differ between pseudonymous and non-pseudonymous articles, so I estimate the following model by OLS, employing propensity-score matching to weight matched pairs and exclude unmatched pairs:

$$car_{i,j,t_0+2,t_0+5} = \beta_0 + \beta_1 car_{i,t_0-1,t_0+1} + \beta_2 pseudo_j + \beta_3 \left(pseudo_j \times car_{i,t_0-1,t_0+1}\right) + \epsilon_{i,j,t+2,t+5}$$

where $car_{i,j,t,\tau}$ is defined above, $pseudo_j$ is 1 if author j is pseudonymous, and $\epsilon_{i,j,t+2,t+5}$ is a random error term. As an additional measure of price reversal, I consider the simple difference between the cumulative abnormal return over the $[t_0+2, t_0+5]$ interval on the cumulative abnormal return over the $[t_0, t_0+1]$ interval:

$$rev_{i,j,t} = car_{i,j,t_0+2,t_0+5} - car_{i,j,t_0-1,t_0+1}$$

This measure increases with the divergence between car_{i,j,t_0+2,t_0+5} and car_{i,j,t_0-1,t_0+1}. For example, if $car_{i,j,t_0-1,t_0+1} = -0.02$ but $car_{i,j,t_0+2,t_0+5} = 0.04$, then $rev_{i,j,t} = 0.06$. Note that this does not incorporate positive reversals. If $car_{i,j,t_0-1,t_0+1} = 0.02$ and $car_{i,j,t_0+2,t_0+5} = -0.04$, then $rev_{i,j,t} = -0.06$. However, if $car_{i,j,t_0-1,t_0+1} = 0.02$ and $car_{i,j,t_0+2,t_0+5} = 0.8$, then $rev_{i,j,t} = 0.06$. $rev_{i,j,t} > 0$ thus corresponds to either a *negative* reversal (i.e., a decline in price followed by a subsequent increase) or a larger increase in price over $[t_0+2, t_0+5]$ than the increase over $[t_0-1, t_0+1]$. This latter case is a kind of positive "correction" in the sense that the increase over $[t_0-1, t_0+1]$ may have been depressed.

I regress $rev_{i,j,t}$ on $pseudo_j$, an indicator for pseudonymous authors, and also compare two additional outcomes: (a) an indicator equal to 1 if $rev_{i,j,t} > 0$ and (b) an indicator equal to 1 if $rev_{i,j,t} > 0.02$. The results are shown in Table 5.

[Table 5]

Table 5 shows that pseudonymous articles are linked to a negative correlation between the post-publication price and the price over the following days: a 1 log-point increase in cumulative abnormal log returns in the window $[t_0-1, t_0+1]$ is followed by a decline of .11 log points of cumulative abnormal returns, on average, in the window $[t_0+2, t_0+5]$, where t is the date of publication. This coefficient estimate is significant at the 5% level. The non-interacted coefficient on car_{i,t_0-1,t_0+1} is positive, indicating that non-pseudonymous articles are not followed by price reversals using the Tetlock (2011) measure.

Similarly, $rev_{i,j,t}$ (the difference between car_{i,j,t_0+2,t_0+5} and car_{i,t_0-1,t_0+1}) is .0080 higher for pseudonymous articles on average, a difference that is significant at the 1% level, whereas $rev_{i,j,t}$ is indistinguishable from zero for non-pseudonymous articles. Similarly, pseudonymous articles are 9.2% more likely to be followed by a positive $rev_{i,j,t}$ of any magnitude (this is obtained by dividing the coefficient .0465 by the intercept term .5047), and nearly 13.2% more likely to be followed by a positive $rev_{i,j,t}$ exceeding 2 log points in magnitude (.0430/.3262 $\approx$ 0.132). This statistical evidence is consistent with the visual pattern displayed in Figure 4.

3.5 Informed Trading in Options Markets

Fox et al. (2018) point out that manipulating markets by making negative misstatements "is actually a special kind of informed trading. Just like a corporate insider who trades on confidential information that he obtained from his employer, the misstatement manipulator makes his purchases on the basis of something that he knows and the market does not: the falsity of the price-depressing misstatement for which he is responsible." (p. 145). A large literature finds that informed traders exploit their informational advantages in options markets, where embedded leverage facilitates larger profit-taking. Chakravarty *et al.* (2004) show that options markets contribute 17% to price discovery. Future stock returns can be predicted both by options volume (Pan & Poteshman, 2006) as well as deviations from put-call parity (Cremers & Weinbaum, 2010). And Mitts & Talley (2018) find that put-option trading volume and open interest rise in the months preceding the disclosure of a cybersecurity breach.

Are these price reversals driven by this kind of manipulative "informed buying" at prices that have been artificially depressed by the publication of a pseudonymous attack article? A measure of bullish or bearish sentiment in options markets is the relative demand for put vs. call options, which has been found to predict informed trading (e.g., Pan & Poteshman (2006)). I examine trading behavior over these windows using individual quotes for equity options that are nearly at-the-money (delta between 0.45 and 0.55) in the OptionMetrics IvyDB for each of the firm-articles in the single-neighbor matched sample ($n = 992,946$).

3.5.1 Open Interest and Volume

It can be difficult to precisely measure informed trading in options markets, so I consider multiple approaches. Prior literature has found that the ratio of demand for put options to call options predicts future stock returns (Pan & Poteshman, 2006), measuring this demand with abnormal open interest, the number of outstanding open put or call contracts, and transaction volume, the number of contracts traded on a given day (Cao *et al.* , 2005; Chakravarty *et al.* , 2004; Jayaraman *et al.* , 2001). Accordingly, I employ a difference-in-difference-in-differences design which compares the over-time difference in the open interest and volume of put vs. call options, as between pseudonymous and non-pseudonymous articles, prior to and following two periods: (1) the date of disclosure (t_0) and (2) the reversal period $[t_0 + 2, t_0 + 5]$.

I begin by plotting over-time trends on the difference in log open interest between pseudonymous and non-pseudonymous articles. Figure 5 plots the average difference in log open interest between pseudonymous and non-pseudonymous articles (i.e., $y_{i,t} = a_{i,t} - n_{i,t}$ where $a_{i,t}$ is log open interest for pseudonymous articles and $n_{i,t}$ is log open interest for non-pseudonymous articles) for put options and call options, separately, after subtracting the average log open interest for calls and puts written on each firm-article in the interval $[t_0 - 9, t_0 + 5]$ (i.e., a fixed effect specification). To examine whether the parallel trends assumption holds, I begin the figure at $t_0 - 9$, where t_0 is the publication date of the article.

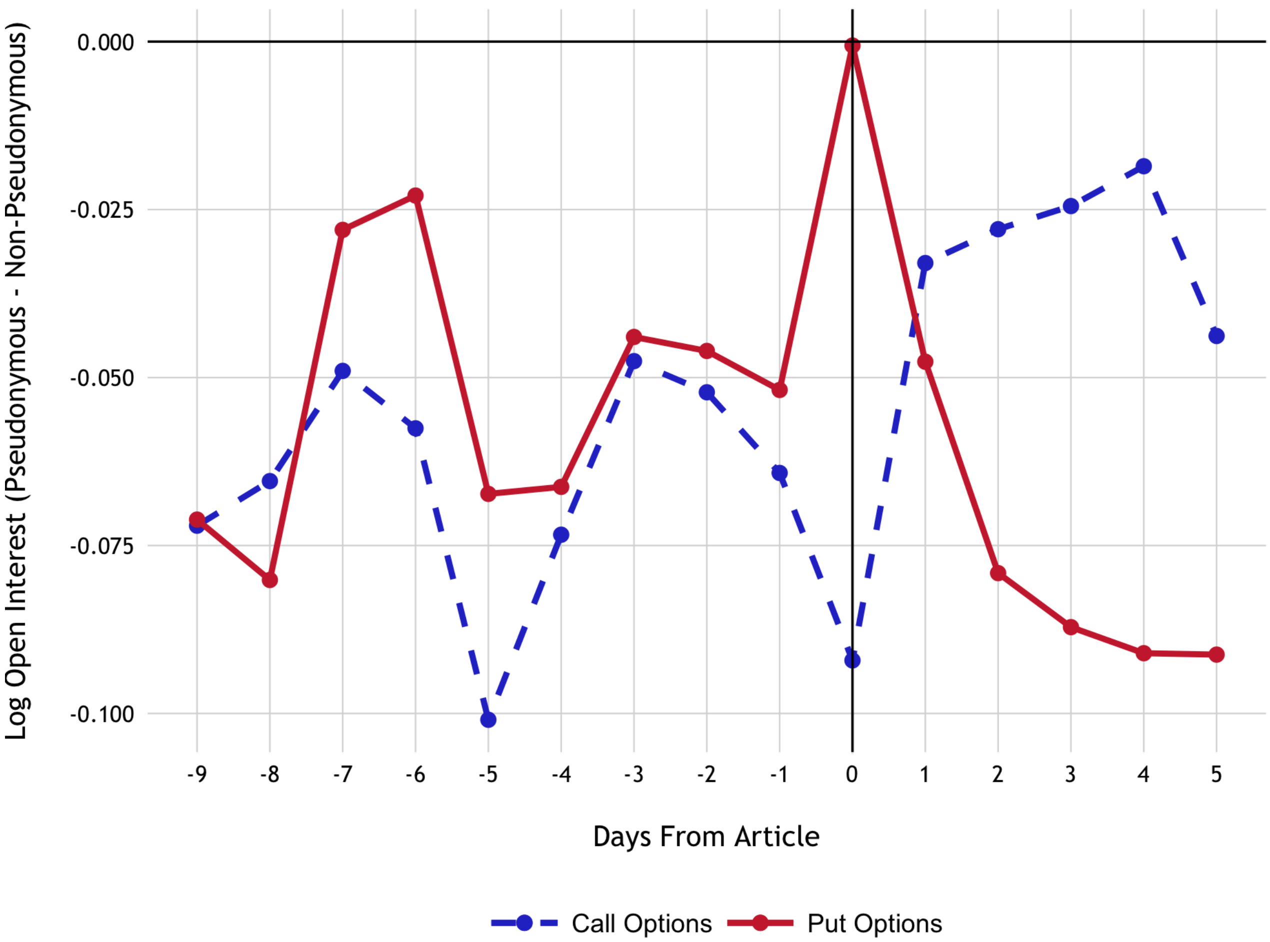
As Figure 5 shows, the trends are largely parallel over the interval $[t_0 - 9, t_0 - 1]$. At t_0, the demand for put options skyrockets, which suggests that some of the price decline on the day of publication may be driven by highly leveraged option trades on the day of publication. This trend flips direction immediately thereafter: from t_0+2 to t_0+5, the period during which prices reverse direction, the demand for call options exceeds the demand for put options. Figure 8 in the Online Appendix shows a similar pattern for log volume.

This evidence is consistent with the manipulative pattern of “informed buying” exploiting the negative reaction to the article, which causes prices to reverse direction during the interval $[t_0 + 2, t_0 + 5]$. To examine this statistically, I estimate two different models by OLS:

$$y_{i,j,t} = \beta_0 + \beta_1 pub_t + \beta_2 call_j + \beta_3 \left(pseudo_j \times pub_t\right) + \beta_4 \left(pseudo_j \times call_j\right)$$
$$+\beta_5 \left(pub_t \times call_j\right) + \beta_6 \left(pseudo_j \times pub_t \times call_j\right) + \alpha_i + \epsilon_{i,j,t}$$

Figure 5: Log Open Interest (Pseudonymous − Non-Pseudonymous Difference)

This figure plots, on the y-axis, the average difference in log open interest between pseudonymous and non-pseudonymous articles, and on the x-axis, the day relative to the date of publication. The blue dashed line plots the average pseudonymous vs. non-pseudonymous difference in open interest for call options and the red solid line plots the average pseudonymous vs. non-pseudonymous difference in open interest for put options. Both are the residuals from a fixed effect specification, i.e., after subtracting the average pseudonymous vs. non-pseudonymous difference in log open interest for calls and puts written on each firm-article in the interval $[t_0 - 9, t_0 + 5]$.



and:

$$y_{i,j,t} = \beta_0 + \beta_1 post_t + \beta_2 call_j + \beta_3 \left(pseudo_j \times post_t\right) + \beta_4 \left(pseudo_j \times call_j\right)$$

$$+\beta_5 \left(post_t \times call_j\right) + \beta_6 \left(pseudo_j \times post_t \times call_j\right) + \alpha_i + \epsilon_{i,j,t}$$

where $y_{i,j,t}$ is log open interest or volume on day t for option j written on the stock of the firm that is the subject of article i; $pseudo_j$ is 1 if the article was published by a pseudonymous author; pub_t is 1 if day $t = t_0$, where t_0 is the publication date of the article; $post_t$ is 1 if day t lies within the correction period of $[t_0 + 2, t_0 + 5]$; $call_j$ is 1 if the option is a call option; α_i is a fixed effect for firm-article i; and $\epsilon_{i,j,t}$ is a random error term. The coefficient of interest is β_6, which captures the difference in open interest or volume between call and put options for pseudonymous articles on the publication day t_0 or during the correction period $[t_0+2, t_0+5]$. Standard errors are clustered by firm-article, and the results are presented in Table 6.

[Table 6]

As Table 6 shows, the triple-difference coefficient β_6 is negative and statistically significant in the publication-day specification ("Pseudonymous × Publication Day × Call Option"). Columns 1 and 2 show that the open interest and volume of a call option written on the target of a pseudonymous article are 7.66 and 8.92 log points lower, respectively, than put options, on the day of publication. Similarly, the triple-difference coefficient β_6 is positive and statistically significant in the correction-period specification ("Pseudonymous × Correction Period × Call Option"). Columns 3 and 4 show that the open interest and volume of a call option written on the target of a pseudonymous article are 7.75 and 6.20 log points higher, respectively, than put options, during the correction period, compared to the day of disclosure.

Can market participants detect manipulative options trading during this period? Table 7 reports the results of estimating this same triple-difference model on observable characteristics of these options: the time-to-expiration, strike price, absolute delta and gamma. Table 7 shows that the triple-different coefficient is statistically indistinguishable from zero on both the day of the article's publication and the reversal day. This indicates that informed trading on the knowledge of a forthcoming manipulative short attack is occurring among options that are observationally similar; pseudonymous authors are not "tipping their hand" by trading in options that expire more quickly or are unusual in some other way.

[Table 7]

3.5.2 Put-Call Parity

Cremers & Weinbaum (2010) find that deviations from put-call parity predict stock returns, suggesting the presence of informed trading. They estimate these deviations by measuring the difference in implied volatility between put and call options with the same strike price and expiration date. Similarly, An *et al.* (2014) find that changes in implied volatility predicts future stock returns. I examine whether deviations from put-call parity predict informed trading during the period of a pseudonymous attack by matching put and call options for a given security on expiration date and strike price, and considering whether implied volatility differs between the matched puts and calls, as in Cremers & Weinbaum (2010).

I estimate the same triple-difference specification as in the prior subsection, but replace the firm-event fixed effect α_i with a fixed effect corresponding to the unique combination of the underlying security, expiration date and strike price. First, I compare the period $[t_0, t_0 + 2]$, when put-call parity should reflect informed trading in the direction of put options, to the baseline period $[t_0 - 9, t_0 - 1]$. Second, I compare the reversal period $[t_0 + 3, t_0 + 5]$, when put-call parity should return to the baseline, to the elevated period $[t_0, t_0 + 2]$. The estimations are lagged by one day in contrast to Table 6 in order to account for options markets updating in response to informed order flow. The results are presented in Table 8.

[Table 8]

As Table 8 shows, implied volatility is higher for put options relative to call options written on the targets of pseudonymous attacks over the window $[t_0, t_0 + 2]$. Similarly, implied volatility is higher for call options relative to put options over the lagged reversal period $[t_0 + 3, t_0 + 5]$. Like Cremers & Weinbaum (2010) and An *et al.* (2014), this deviation from put-call parity indicates the presence of informed trading in options markets over these windows.

Taken together, this evidence is strongly consistent with the kind of misstatement manipulation identified by Fox et al (2018), whereby traders aggressively purchase the stock of target firms while prices are artificially depressed, profiting off the expected price correction. Section 3.7 examines whether this kind of "informed buying" imposes welfare costs in the form of higher bid-ask spreads, as predicted by the literature on informed trading.

3.6 Are Reversals Driven by Provocative Article Content?

One question is whether pseudonymous authors simply induce a price overreaction by posting provocative content. To be sure, it is difficult to explain the purchase of put options prior to the public posting of the article as a mere coincidence, and the analysis in Section 4 below suggests an exploitation of the market's view that the author is trustworthy. This Section provides additional evidence that evocative article content is unlikely to explain these findings.

I generate a document-term matrix for the articles in my dataset following the text analysis literature (Grimmer & Stewart, 2013). First, I extract the raw text of each article and preprocess by removing punctuation and numbers, converting to lowercase, removing "function" words (a, the, etc.), and stemming to merge different grammatical forms. As in Macey & Mitts (2014), I tokenize the preprocessed text by counting the frequencies of individual word, two-word phrases ("bigrams") and three-word phrases ("trigrams"). I discard words and phrases which occur in less than 0.1% of the sample, yielding a document-term matrix of 9,515 articles in the original sample × 107,890 words and phrases. Restricting to the matched sample yields a final document-term matrix of 2,900 articles × 104,495 words and phrases.

I utilize these data in two ways. First, to examine whether pseudonymous articles employ provocative phrases like "fraud", accounting restatements, and so forth, I extract those words and phrases containing the terms "fraud", "restatement" or "legal" (in stemmed form). For each article in my dataset, I count the number of times any of these terms appear, and examine whether this count differs between pseudonymous and real-name articles. The results of this balance test are shown in Table 4, which shows that fraud-related text does not appear more frequently in pseudonymous articles. Indeed, while the difference in means is statistically insignificant, the point estimate is slightly smaller for pseudonymous articles.

While this test is intuitive, it is quite crude, and one might worry that the ex-ante selection of a given set of terms will lead to biased inferences. As an alternative test, I utilize the entire vocabulary of 104,495 words and phrases in the matched sample, and employ logistic regression with elastic-net regularization (Zou & Hastie, 2005) to identify which words and phrases are most predictive of pseudonymous authorship. I choose the regularization parameter λ by ten-fold cross validation, maximizing the area under the ROC curve (the "AUC"). The in-sample accuracy of the model is 0.9000, with a sensitivity (true positive rate) of 0.7568 and a specificity (true negative rate) of 0.9983. The no-information rate is 0.5931.[16]

While the in-sample accuracy is likely biased upward, the goal of this exercise is simply to identify which words and phrases are most predictive of pseudonymous authorship. Among the 104,495 words and phrases in the model, only 620 are assigned nonzero coefficients by the elastic-net model, and only 163 positively predict a pseudonymous article. Table 9 shows the top 100 words and phrases in this group that positively predict a pseudonymous article.

[Table 9]

As Table 9 shows, none of the words and phrases that predict pseudonymous authorship seem particularly provocative or address fraud or similar claims. Rather, they seem to reflect idiosyncratic noise. This evidence suggests that pseudonymous authors manipulate markets by engaging in informed options trading, artificially depressing prices by trading on the advance knowledge of a forthcoming article, and not by promulgating provocative content.

3.7 Bid-Ask Spreads

Glosten & Milgrom (1985) show that the presence of informed trading causes market makers to enlarge bid-ask spreads to compensate for expected trading losses. I examine whether spreads widen for the targets of pseudonymous attacks. Both the acquisition of put options on the day of the attack as well as the accumulation of long positions during the correction period constitute a kind of "informed trading" on the fact that an article does not have fundamental-value implications for the value of the firm. However, market makers are only able to anticipate the latter, because the publication of the article itself comes as a surprise to the market. In anticipation of the accumulation of call options during the correction period (which will be hedged by counterparties opening long positions in the underlying stock), market makers are likely to widen the spread. Figure 4 suggests this occurs at $t_0 + 2$, so a natural starting point is to ask whether bid-ask spreads increase from the day of publication to two days after, when

[16] The no-information rate is the accuracy rate if articles were randomly classified as pseudonymous or not, and is equal to the in-sample proportion of the most frequent class.

informed traders will aggressively begin to purchase the shares of those target firms whose stock prices were artificially depressed by the pseudonymous attack.

I measure bid/ask spreads using daily pricing data reported by the Center for Research on Securities Prices (CRSP). These data are rough approximations, but useful for daily analysis of this kind. Bid/ask spreads are highly persistent, so over-time variations in spreads tend to be multiplicative in nature. A firm with a small spread of \$0.01 is extremely unlikely to see its spread double to \$0.02, even with substantially increased informed trading; but a firm with a spread of \$0.20 could easily see that spread increase to \$0.21. This motivates the following percentage definition of the change in the spread:

$$\Delta spread_{i,t,\tau} = \frac{spread_{i,t}}{spread_{i,\tau}} - 1$$

An alternative normalizes the spread by the price of the underlying stock:

$$\Delta s\hat{pr}ead_{i,t,\tau} = \frac{spread_{i,t}/p_{i,t}}{spread_{i,\tau}/p_{i,\tau}} - 1$$

For the reasons described above, I focus on $\Delta spread_{i,t_0,t_0+2}$ and $\Delta s\hat{pr}ead_{i,t_0,t_0+2}$, i.e., the percentage change in the spread from the day of publication to two days thereafter.

In a competitive market among liquidity providers, market makers will increase the spread commensurately with the risk of informed trading. As of day $t_0 + 2$, market makers observe the extent of the price decline over the interval $[t_0 - 1, t_0 + 1]$, and the analysis in Table 5 indicates that this price decline is a key proxy for the expected reversal. For this reason, I estimate the following model by OLS, employing propensity-score matching at the firm-article level to weight matched pairs and exclude unmatched pairs:

$$\Delta spread_{i,t_0,t+2} = \beta_0 + \beta_1 pseudo_j + \beta_2 car_{i,t_0-1,t_0+1} + \beta_3 \left(pseudo_j \times car_{i,t_0-1,t_0+1}\right) + \epsilon_{i,t,\tau}$$

The key coefficient of interest is β_3, which reflects the percentage-point change in the spread with the cumulative abnormal log return over the interval $[t_0 - 1, t_0 + 1]$. The prediction is that $\beta_3 < 0$, i.e., as car_{i,t_0-1,t_0+1} declines (becomes more negative), the spread increases. I also consider two alternative regressors: an indicator equal to 1 if $car_{i,t_0-1,t_0+1} < 0$, i.e., a negative market reaction to article publication, as well as an indicator equal to 1 if $car_{i,t_0-1,t_0+1} < -0.05$, i.e., a strongly negative market reaction. The results are given in Table 10.

[Table 10]

Column (1) of Table 10 shows that, on average, a 1 log point decrease in the cumulative abnormal log return over the interval $[t_0 - 1, t_0 + 1]$ is linked to an increase of 2.18 to 2.24 percentage points in the bid/ask spread from the day of publication to two days post-publication, when the informed call options trading is expected to commence. Similarly, spreads increase by approximately 43-44 percentage points for pseudonymous articles with a decline in the cumulative abnormal log return over the interval $[t_0 - 1, t_0 + 1]$, and an increase of 50-52 percentage points for $car_{i,t_0-1,t_0+1} < -0.05$, i.e., a strongly negative market reaction. This evidence is consistent with the concerns raised in Fox et al. (2018) that this sort of market manipulation constitutes a form of informed trading that imposes social welfare costs by widening the bid/ask spread.

3.8 Aggregate Trading Losses

What are the aggregate trading losses due to the mispricing caused by the publication of pseudonymous articles? It is important not to confuse these trading losses, which are merely ex post transfers between traders, with the welfare costs of informed trading. Those welfare losses are driven by the reduction in liquidity and increase in the bid-ask spread as a result of pseudonymous market manipulation (Glosten & Putniņš, 2016); here, I simply compute the extent to which trades were executed at an incorrect price ex post.

I consider solely the 1,720 firm-articles written by pseudonymous authors and calculate the aggregate dollar volume of trading on each of the trading days from $[t_0, t_0 + 4]$, excluding $t_0 + 5$ because that is used to calculate the counterfactual price. I then calculate the counterfactual dollar volume by multiplying the number of shares that were traded for each firm by the price of the firm on $t_0 + 5$. This is the price that sellers would have received in the absence of any price distortion, i.e., if the shares had been sold at their price on day $t_0 + 5$. To calculate *net* mispricing, I subtract the actual dollar volume from the counterfactual dollar volume, which measures the price sellers would have received if the counterfactual price at t_0+5 had prevailed over those days. The price at t_0+5 may be greater than or less than the price on the days $[t_0, t_0+4]$, but is higher on average. This calculation is given in Table 11.

[Table 11]

As Table 11 shows, sellers would have received a total of $20.1 billion more during the interval $[t_0, t_0+4]$ if trades had been executed at the price on t_0+5. In Section 5, I discuss whether the price distortion induced by pseudonymous attacks may give rise to liability under Section 9(a) of the Securities Exchange Act of 1934 and Rule 10b-5.

4 Why Listen to Pseudonymous Authors?

4.1 Reputation Theory and Pseudonymous Manipulation

Why does pseudonymous market manipulation persist in an efficient market with sophisticated investors who have billions of dollars at stake? A useful starting point is the canonical model of market manipulation articulated in Benabou & Laroque (1992), which I briefly review. In Benabou & Laroque (1992), there are two states of nature (good and bad), which are equally likely, and one risky asset which pays $1 in the good state and nothing in the bad state. The key to their model is the assumption that the author observes a private signal s which is imperfectly correlated with the asset payoff:

$$Pr\,(\text{ good state } \mid s=1) = p$$

with $p \in [1/2, 1]$. The author chooses to send a message $m \in 0, 1$ (i.e., "good news" or "bad news"), and the probability that the author is truthful is given by $q \in [0, 1]$:

$$Pr\,(m=1 \mid \text{good state }) = Pr\,(m=0 \mid \text{bad state }) = q$$

$$Pr\,(m=0 \mid \text{good state }) = Pr\,(m=1 \mid \text{bad state }) = 1-q$$

Rational agents price the asset via Bayesian updating, adjusting their prior belief of $1/2$ by the message they receive. The key question is how much *credibility* they ascribe to the message, which formally turns on q. Letting ρ denote agents' prior belief that $q = 1$, Benabou & Laroque (1992) show that the posterior belief as to the quality of the author's information p, denoted $\pi \in [p, 1-p]$, conditional on observing a message m, is given by:

$$\pi(\rho|p, q) = \rho p + (1-\rho)\left[pq + (1-p)(1-q)\right]$$

i.e., with probability ρ the message is credible ($q = 1$) and thus the asset pays off with probability p. With probability $1-\rho$, the message is not credible, which means that the probability that the asset pays off depends on q: with probability q, the asset pays off p, and with probability $1-q$, the asset pays off $1-p$ instead of p. This simplifies to the following:

$$\beta = Pr(\mid \text{good state } m = \{1, -1\}) = \frac{1}{2} + m\left(\pi(\rho) - \frac{1}{2}\right)$$

where β denotes the market's posterior as to the asset payoff, conditional on observing m. The market's valuation turns on m and q, i.e., the message sent by the author as well as his or her credibility, which is what facilitates manipulation. Benabou & Laroque (1992) show that in a repeated setting, agents' update their beliefs as follows. If the time-t message is correct, i.e., $m_t = 1$ and the good state realizes or $m_t = 0$ and the bad state realizes, Bayes' Rule yields:

$$\rho_{t+1} = \frac{p\rho_t}{\pi(\rho_t)}$$

and if the time-t message is incorrect:

$$\rho_{t+1} = \frac{(1-p)\rho_t}{1-\pi(\rho_t)}$$

It is immediately apparent that if $p = 1$, i.e., the author's private signal perfectly predicts the asset payoff, a single incorrect message leads to $\rho_{t+1} = 0$ and the market disregards future messages. This is a simple formalization of the point in Putniņš (2012) that "if market participants are able to deduce that false information originated from a manipulator, the manipulator will quickly be discredited and the manipulation strategy will cease to be profitable." In light of this result, Benabou & Laroque (1992) focus on $p < 1$, leading to an updated prior $\rho_{t+1} \in (0, 1)$ and allowing the manipulation to continue even after an incorrect message.

But this result turns on the market's ability to continue to attribute the message at time $t+1$ to the same author. It is easy to see that if an author is able to "reset" the market's prior by setting $\rho_{t+1} = \rho_0$, where $\rho_0 > 0$ is the initial prior on the author's credibility, then even if $p = 1$, the market will respond to a manipulative message going forward. Pseudonymity serves this function by allowing those authors who lack credibility to reset the market's prior as to their credibility, even if $p = 1$ (or p is very close to 1). The net effect is a kind of pooling equilibrium: absent a history of manipulation, market participants cannot separate a manipulative from non-manipulative article when a new pseudonym emerges. As discussed previously, Table 7 shows there is no difference in observable options characteristics which might signal to market participants that manipulation is taking place.

That said, a puzzle remains: why is pseudonymity itself tolerated by markets? Why not set $\rho_{t+1} = 0$ for a pseudonymous author? The classical "unraveling" result in the disclosure literature shows that a seller who possesses private information that a good is high-quality has an incentive to fully disclose this information so as to induce buyers to pay for the quality; absent disclosure, buyers will assume that the seller has something to hide (Grossman, 1981). Pseudonymity invites this kind of adverse inference: market participants might rationally conclude that authors with truthful information should have no trouble risking their reputation by making claims using their real name; after all, if the information is true, no harm to their reputation will result. The use of pseudonymity suggests that an author has something to hide. Rational investors should simply ignore pseudonymous articles, inferring that if a claim is truthful, it will be made by an author using their real name.

But that kind of inference breaks down when authors may have a reason to prefer pseudonymity other than false information. Indeed, both Figure 4 and Table 5 show that trading on pseudonymous attacks is *profitable* on average — just not as profitable as it would be absent the price reversals documented in this article. Consistent with this finding, Seeking Alpha justifies its pseudonymity policy by pointing out that "regulations at their workplace or other factors" prevent "some contributors [from] revealing their real names. In addition, many well-known, veteran stock market bloggers (some of the finest, in fact) write under a pseudonym."[17] The willingness of market participants to sell the stock of targets of pseudonymous attacks can be rationalized by pointing to the ambiguity underlying the use of a pen name: authors may prefer pseudonymity precisely because they are conveying truthful information and fear the adverse consequences that may result from being identified as the author of such truthful analysis.

In addition to workplace prohibitions on social media commentary, authors may also fear litigation risk. Attackers may worry that target firms will pursue defamation or securities fraud claims if the publication of an attack piece leads to a decline in the price of the stock. Even if the author can fully establish the truth of every claim made in the piece, doing so would involve protracted, time-consuming litigation, which imposes nontrivial costs. Pseudonymity allows these authors to make damaging but truthful claims without worrying that target firms will bring an unfounded lawsuit that is costly to defend against.

Yet another example is legal uncertainty: the precise contours of securities fraud liability are not always clear. Consider, for example, the Supreme Court's 2015 decision in *Omnicare, Inc. v. Laborers District Council Construction Industry Pension Fund*, which overturned the Sixth Circuit's holding that statements of opinion that ultimately turn out to be incorrect may constitute an "untrue statement of a material fact" under Section 11 of the Securities Act of 1933. Instead, the Court held that opinions may constitute misstatements when they are not sincerely held. This kind of uncertainty in the doctrinal landscape may lead authors to prefer pseudonymity to make it more difficult to be held accountable for violating a legal rule whose interpretation is shifting and subject to judicial clarifications ex post.

For these reasons, market participants may be hesitant to conclude that pseudonymity necessarily implies a lack of credibility. But by allowing authors to effectively switch names, pseudonymity undermines the effectiveness of the reputation mechanism envisioned in Benabou & Laroque (1992). Their model seems to implicitly assume a lack of pseudonymity: "[i]f the insiders' information was perfect, one could easily tell ex post whether or not they had been

[17] https://seekingalpha.com/page/policy_pseudonymous_contributors

truthful. **In this case they could lie at most once, and sanctioning fraud would eliminate the problem**" (p. 924, emphasis added). They further argue that "in reality even private information is not fully reliable, so that the possibility of honest mistakes makes it very difficult to establish fraud conclusively." Yet the mere possibility of honest mistakes is not a roadblock to establishing fraud: there is often evidence as to whether a given misstatement was driven by deceptive intent or not. On the other hand, it is difficult to sanction pseudonymous authors for fraud measured by ex-post price reversals, as discussed in Section 5.

This theoretical framework yields several testable predictions. First, pseudonymous authors should focus informed trading on those times when they are trustworthy, i.e., $\rho_{t+1} > 0$ in the model of Benabou & Laroque (1992), such as when they have no history ($\rho_{t+1} = \rho_0$) or their history has had few mistakes so $\rho_{t+1} > 0$. Second, pseudonymous authors should "disappear" after the market realizes they have been misled, so that they can switch to a new identity. Finally, identity-switching by pseudonymous authors should leave traces of underlying authorship, which are detectable using techniques of linguistic stylometry.

4.2 Pseudonymity and Trustworthy Trading

I test the prediction that pseudonymous authors will engage in informed trading when they are "trustworthy," such that $\rho_{t+1} \neq 0$. For an anecdotal example of this kind of behavior, consider again the SkyTides-Insulet case. Figure 6 was taken from SkyTides' website, and shows the short seller's history of success prior to Insulet. It is clear that the market was justified in listening to SkyTides, as the pseudonymous author had established a trustworthy track record prior to attacking Insulet while purchasing put options prior to publication of the article.

I systematically test the hypothesis that pseudonymous authors exploit the market's trust by defining, for each firm-article, "trustworthy" as the absence of any reputational history (i.e., the author's first article on Seeking Alpha) or a prior history of non-reversals. I define

Figure 6: SkyTides History of Trustworthy Trading

This figure shows a screenshot from SkyTides' website, which shows that the pseudonymous attacker had accumulated a history of successful non-reversals prior to the Insulet case.

Company	Ticker	Rating	Date Report Released	$ Stock Price at Report Release	$ Stock Low Price Since Report Release	Stock Price % Decline	Officer or Board Departures
Insulet Corporation	PODD	Sell Short	11/15/16	$ 34.63	$ 32.12	-7.2%	None
Vocera Communications, Inc.	VCRA	Sell Short	4/5/16	$ 12.23	$ 10.46	-14.5%	Board Members
Conformis, Inc.	CFMS	Sell Short	1/19/16	$ 13.39	$ 3.89	-70.9%	CEO
Tantech Holdings Ltd	TANH	Sell Short	9/24/15	$ 21.39	$ 1.00	-95.3%	None
HII Technologies, Inc.	HIIT	Sell Short	12/30/13	$ 0.52	$ 0.00	-99.8%	CEO, CFO
Green Automotive Company	GACR	Sell Short	12/19/13	$ 0.25	$ 0.00	-99.1%	CEO, CFO

the following variable for the article published by author j about firm i at time t:

$$trustworthy_{i,j,t} = \begin{cases} 1 \text{ if } t = 0 \text{ or } \sum_{\tau=0}^{t} rev_{k,\tau} < 0 \ \forall k \in J_{k,t} \\ 0 \text{ otherwise} \end{cases}$$

where $J_{k,t}$ is the set of articles written by author j prior to time t, with each article indexed by k. I estimate the same regression model as in Section 3.4, comparing the sample where $trustworthy_{i,j,t} = 1$ to $trustworthy_{i,j,t} = 0$. The results are presented in Table 12.

[Table 12]

Table 12 shows that negative reversals occur when authors are perceived as trustworthy by the market. I examine whether the open interest findings in Table 6 are driven by trading at trustworthy times.[18] I estimate the same model as in Section 3.5, comparing the sample where $trustworthy_{i,j,t} = 1$ to $trustworthy_{i,j,t} = 0$. The results are given in Table 13.

[Table 13]

As Table 13 shows, informed trading in options markets is occurring when pseudonymous authors are perceived as trustworthy by the market. It is possible that the lack of statistical

[18] Volume yields qualitatively similar but noisier estimates, as expected with many days having no volume.

significance in the non-trustworthy subsample is driven by insufficient power, but notice that these samples have $n = 226,764$ and $n = 86,931$. Taken together, this evidence is consistent with the theoretical prediction in Benabou & Laroque (1992) that informed trading will be concentrated in cases where the market perceives an author as trustworthy.

4.3 Pseudonymity and Disappearing Authors

A second implication of this theoretical framework is that pseudonymous authors should switch identities once the market realizes they promulgating misleading articles. In this Section, I examine whether pseudonymous authors are more likely to "disappear" after it is apparent that the market is no longer listening to what they have to say. I test three distinct propositions.

First, I examine whether pseudonymous authors are more likely to "disappear," i.e., whether a given article is likely to be the last one written by an author. I estimate the following regression on the matched sample:

$$last_{i,j,t} = \beta_0 + \beta_1 pseudo_j + \epsilon_{i,j,t}$$

where $last_{i,t}$ is 1 if article i written at time t is the last one by author j, $pseudo_j$ is 1 if author j is pseudonymous, and $\epsilon_{i,j,t}$ is a random error term.

Second, I test whether the market response to an article is linked to the presence or absence of prior reversals. For each author I derive the mean of prior negative reversals, which is based on the same metric used in the prior section to determine trustworthy periods:

$$prior_{i,j,t} = \frac{1}{N_{J_{k,t}}} \sum_{\tau=0}^{t} rev_{k,\tau} < 0 \ \forall k \in J_{k,t}$$

where $J_{k,t}$ is the set of articles written by author j prior to time t, with each article indexed by k, and $N_{J_{k,t}}$ denotes the length of $J_{k,t}$. I define the "market response" to an article as $|car_{i,j,t_0-1,t_0+1}|$, i.e., the absolute abnormal return over the interval $[t_0 - 1, t_0 + 1]$,[19] and

[19]The closer $|car_{i,j,t_0-1,t_0+1}|$ is to zero, the less stock prices changed in response to the publication of the article.

estimate the following regression by OLS on the matched sample:

$$|car_{i,j,t_0-1,t_0+1}| = \beta_0 + \beta_1 prior_{i,j,t} + \epsilon_{i,j,t}$$

where $last_{i,j,t}$ is defined above and $\epsilon_{i,j,t}$ is a random error term.

Finally, I link the two prior tests together and consider whether pseudonymous authors are more likely to "disappear" when the market has ceased to respond to the publication of an article. I estimate the following regression by OLS on the matched sample:

$$last_{i,j,t} = \beta_0 + \beta_1 pseudo_j + \beta_2 |car_{i,j,t_0-1,t_0+1}| + \beta_3 \left(|car_{i,j,t_0-1,t_0+1}| \times pseudo_j\right) + \epsilon_{i,j,t}$$

In addition, I consider an alternative specification where I define a variable which reflects a lack of credibility for a pseudonymous author as follows:

$$low\ credibility\ pseudo_{i,j,t} = \begin{cases} 1 \text{ if } |car_{i,j,t_0-1,t_0+1}| < 0.01 \text{ and } prior_{i,j,t} > 0.05 \text{ and } pseudo_j = 1 \\ 0 \text{ otherwise} \end{cases}$$

The results of these estimations are given in Table 14.

[Table 14]

Column (1) of Table 14 shows that the last article for an author is more likely to be written by a pseudonymous author than a real-name author, which is consistent with pseudonymous authors switching identities. Column (2) shows that the market response to a given article decreases as the author accumulates a history of negative reversals. Column (3) shows that the last article for an author is especially likely to have been written by a pseudonymous author with a low market response—the negative coefficient indicates that for pseudonymous authors, the probability of the last article increases as the market response to the article decreases. Column (4) shows that low-credibility articles by pseudonymous authors are extremely likely to be the last articles written by these authors. Taken together, this evidence is consistent with the theoretical prediction that pseudonymous authors disappear when they lose credibility.[20]

[20]In unreported estimations, I re-ran the analysis in column (4) on non-pseudonymous articles, and find a similar

4.4 Detecting Identity-Switching with Linguistic Stylometry

If pseudonymous authors switch identities after losing credibility with the market, it might be possible to detect the adoption of a new identity using methods of authorship attribution from the field of linguistic stylometry.[21] Stylometry is a technique to identify subtle aspects of an author's writing style that appear throughout documents he or she has written. Stylometric methods have been applied for decades to shed light on the authorship of historical and religious texts like the Federalist Papers (Mosteller & Wallace, 1964) and Book of Mormon (Holmes, 1992), as well as forensic applications (Iqbal *et al.* , 2010).

Stylometry bears many similarities to prior applications of the analysis of textual data (Grimmer & Stewart, 2013; Macey & Mitts, 2014; Varian, 2014). The key difference is that stylometric prediction exploits *non-content* features that are intended to capture the author's writing style rather than the subject matter of the document. Typical techniques in textual analysis such as "bag-of-words" features and discarding so-called function words (like "the", "a", etc.) will lead to effective prediction of a document's substantive content — e.g., a similar company or subject matter in another article. However, the goal of a stylometric analysis is to identify authorship *regardless* of the underlying content he or she has produced. That necessitates using a different set of predictors which are unrelated to document content.

Traditionally, stylometric analysis was performed on a small number of documents, with the goal of extracting a tremendous amount of nuanced detail from the author's available writing. The computational demands of the traditional approach are not well-suited to classifying thousands of documents produced online, such as the Seeking Alpha articles in my dataset. Narayanan *et al.* (2012) solve this problem by identifying predictors that can be used for rapid stylometric analysis on a large scale. I adopt the stylometric features identified in Narayanan *et al.* (2012) to predict authorship for the articles in my dataset.[22]

result—while columns (1)-(3) are significantly different for pseudonymous articles, this suggests that when an author has truly lost credibility, they cease posting regardless of whether they are pseudonymous or not.

[21]Seeking Alpha may employ technology to prevent identity-switching, but it is unclear whether these methods are immune to sophisticated techniques like switching IP addresses or routing over Tor.

[22]These include the number of characters; Yule's K; the frequency of hapax legomena, dis legomena, and so forth; frequency of words with upper case, all lower case, only first letter upper case, camel case (CamelCase); frequency of words with 1-20 characters; frequency of a-z; frequency of 0-9; frequency of punctuation and other special characters; and frequency of function words.

The empirical design is simple. For each article in my dataset, I compare its stylometric features to those of articles written by "former authors," i.e., those authors who had published their final article before that date.[23] Just as before, I infer when an author has published their last article by observing ex post (as of May 2018) what the date of their last article was. The benefit of hindsight should pose no problem: nothing here assumes that market participants were contemporaneously aware of whether a given author would cease to publish.

I compare the stylometric similarity between a given article and this candidate set of articles by calculating the pairwise cosine similarity, defined as follows between article i and k:

$$similarity_{i,k} = \frac{\mathbf{x}_i \cdot \mathbf{x}_k}{||\mathbf{x}_i||||\mathbf{x}_k||}$$

where $\mathbf{x}_i$ denotes the stylometric feature vector for article i. I then derive an article-level similarity measure as the average of these pairwise similarities:

$$similarity_i = \frac{1}{N_k} \sum_{k=1}^{N_k} similarity_{i,k}$$

where N_k denotes the number of articles written by former authors as of the publication date of article i. I calculate this measure for every article in my dataset, not only the matched sample, but then just as before, I perform my analysis on the matched sample, regressing $similarity_i$ on an indicator equal to 1 if article i's author was pseudonymous. Moreover, I examine whether pseudonymous authors hide their identities when first appearing, but eventually revert to their true writing style. I estimate the following specification:

$$similarity_i = \beta_0 + \beta_1 pseudo_j + \beta_2 authorcount_{i,j} + \beta_3 \left(pseudo_j \times authorcount_{i,j}\right) + \epsilon_i$$

where $authorcount_{i,j}$ is the ordinal number of article i for author j, i.e., 1 corresponds to the first article, 2 corresponds to the second article, and so forth. Similarly, I consider specifications where $authorcount_{i,j}$ is replaced by an indicator equal to 1 if $authorcount_{i,j} < 5$ or

[23] Clearly, the author of the current article cannot be included because the final date for that author is, by definition, on or after the publication date of the current article.

$authorcount_{i,j} < 10$. The results are presented in Table 15.

[Table 15]

As Table 15 shows, pseudonymous authors are unconditionally more similar to former authors, but this difference is only marginally significant at the 10% level. However, there is a striking heterogeneity: as pseudonymous authors write more articles, they become far more similar to authors who had published their final article.[24] The specifications with indicators for the author's 5th or 10th article show that pseudonymous authors are significantly more likely to write in a style similar to former authors after the first few articles.

These stylometric findings shed light on an additional mechanism by which pseudonymous authors persuade the market to view their initial article as trustworthy. By adopting a writing style distinct from prior identities, new pseudonymous authors reinforce investors' belief that they are unrelated to prior manipulators. However, after establishing credibility under a new identity, pseudonymous authors no longer find it necessary to write unnaturally. Instead, they can exploit their existing credibility to engage in profitable market manipulation.

5 Pseudonymity and Securities Law

One of the challenges with addressing the sort of market manipulation documented here is that pseudonymous attacks are not easily captured by either the anti-manipulation or anti-fraud provisions of the securities laws, as discussed in the following Sections.

5.1 Market Manipulation

Section 9 of the 1934 Act prohibits a specific set of manipulative trading practices, such as wash sales and matched orders, which reflect artificial trading activity designed to mislead investors as to the underlying interest in the security.[25] Section 9(a)(2) also prohibits "effect[ing] . . . a series of transactions in any security . . . creating actual or apparent active trading in such

[24] This calculation excludes the current author, so the result is not driven by a mechanical correlation.
[25] *See, e.g.*, Ernst & Ernst v. Hochfelder, 425 U.S. 185, 205 n.25 (1976); Edward J. Mawod & Co. v. S. E. C., 591 F.2d 588 (10th Cir. 1979)

security, or raising or depressing the price of such security, for the purpose of inducing the purchase or sale of such security by others."[26] Similarly, subpart (c) of Rule 10b-5 provides that it is unlawful "To engage in any act, practice, or course of business which operates or would operate as a fraud or deceit upon any person, in connection with the purchase or sale of any security."[27]. Market manipulation can run afoul of both Section 9(a)(2) and Rule 10b-5(c).

The Supreme Court has defined market manipulation as "intentional or willful conduct designed to deceive or defraud investors by controlling or artificially affecting the price of securities."[28] Similarly, the Second Circuit has identified the "gravamen of manipulation" as "deception of investors into believing that prices at which they purchase and sell securities are determined by the natural interplay of supply and demand, not rigged by manipulators."[29] Nonetheless, the most challenging question is when trading on the open market is "artificial," absent specific behavior like wash sales or matched orders which are clearly manipulative. [30]

Lower courts disagree on whether open-market trading can violate Section 9(a).[31] It is clear that a large volume of short sales is not per se manipulative because a short seller's bearish view does not, on its own, "mislead investors by artificially affecting market activity."[32] In *SEC v. Masri*, the U.S. District Court for the Southern District of New York held that "an investor conducts an open-market transaction with the intent of artificially affecting the price of the security, and not for any legitimate economic reason, it can constitute market manipulation."[33]

In circuits that decline to follow *Masri*, it may be harder to establish that pseudonymous attacks violate Section 9(a).[34] And while *Masri* represents an expansive view of market ma-

[26] 15 U.S.C. 78i (1970)

[27] 17 CFR 240.10b5

[28] Ernst & Ernst v. Hochfelder, 425 U.S. 185, 198 (1976).

[29] Gurary v. Winehouse, 190 F.3d 37, 45 (2d Cir.1999).

[30] The Seventh Circuit has held that plaintiff bringing a claim under Section 9(a) must establish: "(1) a series of transactions in a security created actual or apparent trading in that security or raised or depressed the market price of that security; (2) the transactions were carried out with scienter; (3) the purpose of the transactions was to induce the security's sale or purchase by others; (4) the plaintiffs relied on the transactions; and (5) the transactions affected the plaintiff's purchase or selling price." AnchorBank, FSB v. Hofer, 649 F.3d 610, 61617 (7th Cir. 2011) But these tests do not shed light on what constitutes "artificial" trading.

[31] *Compare, e.g.*, GFL Advantage Fund, Ltd. v. Colkitt, 272 F.3d 189 (3d Cir. 2001) to SEC v. Masri, 523 F. Supp. 2d 361 (S.D.N.Y. 2007).

[32] In re Scattered Corp. Sec. Litig., 844 F. Supp. 416, 420 (N.D. Ill. 1994) (citing Santa Fe Indus., Inc. v. Green, 430 U.S. 462, 476 (1977)).

[33] 523 F. Supp. 2d at 372.

[34] *See, e.g.*, ScripsAmerica, Inc. v. Ironridge Glob. LLC, 56 F. Supp. 3d 1121, 1162 (C.D. Cal. 2014) ("[Masri] is

nipulation, plaintiffs must still establish intent to artificially affect the price of the security. This presents an evidentiary challenge. Suppose the SEC were able to identify a single individual who authored the attack, purchased put options immediately prior to publication, and similarly bought call options a day or two after the price decline. Even that sort of well-timed trading exploiting these price changes may be insufficient. The author-trader could argue that he or she opened a position consistent with the view articulated in the article that the firm was overvalued, but upon recognizing that investors had overreacted (i.e., the price had declined too far), purchased the stock at that point to take advantage of its return to its fundamental value. Aggressive options trading accompanying the publication of the article may be insufficient, *absent more*, establish intent to artificially depress the price of the security.

One way to establish that sort of intent is to show that the options trading was not merely directionally correct but actually had the effect of distorting the price at the time of publication. In the words of the Second Circuit in *ATSI*, "short selling — even in high volumes — is not, by itself, manipulative. . . . To be actionable as a manipulative act, short selling must be willfully combined with something more **to create a false impression of how market participants value a security**."[35] When options trading is so intense and well-timed with the release of an attack article that the trading gives "a false impression of how market participants value a security," this sort of conduct may constitute market manipulation.[36]

Establishing this sort of manipulative intent requires detailed evidence as to the nature of this sort of trading behavior over time, especially in relation to publication of the attack article. Pseudonymity can make it more difficult for enforcement authorities to determine the identity of the author, as they must subpoena account records, trace IP addresses, and link the author to his or her trading activity. Technologies like Tor can make it harder to "connect the dots." And because the enforcement action will require affirmative evidence of manipulative intent, pseudonymity makes it more likely that regulators will allocate limited resources elsewhere, unless direct evidence of manipulative trading can be brought to bear.

an out-of-circuit district court case that is not binding on the court.").

[35]ATSI Commc'ns, Inc. v. Shaar Fund, Ltd., 493 F.3d 87, 101 (2d Cir. 2007).

[36]*See* Sharette v. Credit Suisse Int'l, 127 F. Supp. 3d 60, 82 (S.D.N.Y. 2015) ("[O]pen-market transactions that are not, in and of themselves, manipulative or illegal, may constitute manipulative activity within the meaning of Section 10(b) when coupled with manipulative intent.")

5.2 Misstatement Fraud

An alternative basis for holding pseudonymous authors accountable is the general anti-fraud rule under subpart (b) of Rule 10b-5, which prohibits "mak[ing] any untrue statement of a material fact."[37] To be sure, the procedural hurdles to bringing a misstatement claim are higher. The Second Circuit has held that misrepresentation claims that do not involve manipulative trading may not be brought under subparts (a) or (c) of Rule 10b-5, and are thus subject to the heightened pleading requirements of the Private Securities Litigation Reform Act (the "PSLRA").[38] The PSLRA mandates that plaintiffs bringing an action for misstatement fraud under subpart (b) of Rule 10b-5 must "specify each statement alleged to have been misleading, the reason or reasons why the statement is misleading, and, if an allegation regarding the statement or omission is made on information and belief, the complaint shall state with particularity all facts on which that belief is formed."[39] While the heightened pleading standard under the PSLRA only applies to private plaintiffs, the SEC must still "stat[e] with particularity the circumstances constituting fraud" under Rule 9(b) of the FRCP.

There are two challenges with applying Rule 10b-5(b) to pseudonymous attacks. The first is that authors are often careful to avoid making factual claims directly. Consider SkyTides' attack on Insulet. The statement that Insulet "allegedly directed employees to bribe physicians" is not a claim that Insulet in fact directed those employees in that manner; rather, it is simply a report of allegations by another. The report refers to a lawsuit filed by one of Insulet's former employees who alleged that the firm's CEO instructed him to "bury" any data that would make the firm look bad. SkyTides could argue that merely reporting a claim made by another is not the same as making the claim. And to the extent that a pseudonymous attacker is merely expressing a negative opinion about a firm, it will be difficult for the SEC

[37] 17 C.F.R. 240.10b-5(b).

[38] Lentell v. Merrill Lynch & Co., Inc., 396 F.3d 161, 177 (2d Cir.2005) ("[W]here the sole basis for such claims is alleged misrepresentations or omissions, plaintiffs have not made out a market manipulation claim under Rule 10b5(a) and (c), and remain subject to the heightened pleading requirements of the PSLRA.) (citing Schnell v. Conseco, Inc., 43 F.Supp.2d 438, 44748 (S.D.N.Y.1999)); accord In re Alstom SA, 406 F.Supp.2d 433, 475 (S.D.N.Y.2005) ([I]t is possible for liability to arise under both subsection (b) and subsections (a) and (c) of Rule 10b5 out of the same set of facts, where the plaintiffs allege both that the defendants made misrepresentations in violations of Rule 10b5(b), as well as that the defendants undertook a deceptive scheme or course of conduct that went beyond the misrepresentations.).

[39] 15 U.S.C. 78u4(b)(1).

to establish that such a view was not "sincerely held."[40]

A more promising possibility is that the pseudonymous attacker lacks a factual basis for the opinions expressed in the piece, such that the attack is a"a misstatement of the psychological fact of the speaker's belief in what he says."[41] The Court in *Omnicare* referred to this as an "embedded statement[] of fact,"[42] that is, by authoring a short attack on a firm, the author is implying that they have some reason to believe in the truthfulness of the underlying claim. The complete absence of any factual basis whatsoever could thus give rise to liability, even if the author's statement consists solely of opinion.

To be sure, litigants have not been very successful on bringing similar claims under defamation law against Seeking Alpha. In *Nanoviricides, Inc. v. Seeking Alpha, Inc.*, the New York Supreme Court for New York County rejected a discovery motion brought against Seeking Alpha to reveal the identity of a pseudonymous author "The Pump Terminator," concluding that "the alleged defamatory statements identified in the petition constitute protected opinion and are not actionable as a matter of law."[43] That said, the decision in *Nanoviricides* was premised on defamation law, not securities fraud, and the court did not undertake an analysis of embedded facts under *Omnicare*.

The second challenge is that establishing a violation of Rule 10b-5 requires proof that the defendant acted with scienter,[44] which in most jurisdictions may be satisfied either by showing an intent to defraud or recklessness.[45] As with market manipulation, it will be difficult to prove an affirmative intent to defraud. And rarely will any factual misrepresention rise to the level required to establish recklessness, i.e., "a danger of misleading buyers or sellers that is either known to the defendant or is so obvious that the actor must have been aware of it."[46]

[40]Omnicare, Inc. v. Laborers Dist. Council Const. Industry Pension Fund, 135 S.Ct. 1318 (2015)).

[41]Virginia Bankshares, Inc. v. Sandberg, 501 U.S. 1083, 1095 (1991). While *Virginia Bankshares* was decided under section 14(a) of the Exchange Act, the *Omnicare* court applied its reasoning to misstatement claims under Section 11 and thus by extension to Rule 10b-5.*Omnicare*, 135 S.Ct. at 1326 n.2.

[42]*Id.* at 1327. (quoting Virginia Bankshares, 501 U.S. at 1109 (Scalia, J., concurring in part and concurring in judgment) ("a statement can sometimes be most fairly read as affirming separately both the fact of the [speaker's] opinion and the accuracy of the facts given to support or explain it.").

[43]2014 WL 2930753 (N.Y. Sup. Ct., Jun. 26, 2014.)

[44]Aaron v. SEC, 446 U.S. 680, 691 (1980)

[45]*See, e.g.*, IIT v. Cornfeld, 619 F.2d 909, 923 (2d Cir.1980); S.E.C. v. Platforms Wireless Int'l Corp., 617 F.3d 1072, 1093 (9th Cir. 2010).

[46]S.E.C. v. Platforms Wireless Int'l Corp., 617 F.3d 1072, 1094 (9th Cir. 2010) (quoting Hollinger v. Titan Capital Corp., 914 F.2d 1564, 156869 (9th Cir. 1990)

The difficulty with bringing an enforcement action on the basis of pseudonymous online postings was highlighted by the SEC's 2007 investigation of John Mackey, CEO of Whole Foods.[47] From 1999-2007, Mr. Mackey posted positive comments on online forums using a fake pseudonyms, including a reference to a company that Whole Foods was considering acquiring. The SEC ultimately closed its investigation into Mr. Mackey's postings without commencing an enforcement action.[48] While the SEC did not state the reasons for dropping its investigation, and Mr. Mackey's case implicated Regulation FD because he was a corporate insider, the difficulties highlighted here likely played a role. The mere use of a pseudonym is unlikely to render authentic expressions of opinion (whether positive or negative) subject to misstatement liability under Rule 10b-5, absent a false factual claim or clear deception.

That said, the recent enforcement action in the matter of *Lidingo Holdings* suggests that the SEC may be willing to take action when such deception can be clearly established. Lidingo Holdings hired writers to publish articles on investment websites like Seeking Alpha under pseudonymous names like "VFC's Stock House." These articles sought to portray certain publicly traded firms in a positive light — namely, those clients of Lidingo who paid for this sort of stock promotion service. The authors did not disclose that they were being compensated by Lidingo for this purpose, nor did they disclose that Lidingo was compensated by the target firms of the articles seeking promotion of their stock.

The SEC brought suit claiming a violation of Section 17(b) of the Securities Act of 1933, which prohibits describing a security for consideration "without fully disclosing the receipt . . . of such consideration and the amount thereof." The SEC also alleged a violation of Rule 10b-5 for the misleading disclosure.[49] The *Lidingo* action shows that when specific factual misstatements or omissions can be established — such as misleading compensation disclosure — the SEC is more willing pursue an anti-fraud enforcement action under Rule 10b-5. But that is a fairly narrow basis on which to impose liability against pseudonymous attackers.

[47] *See, e.g.*, Kara Scannell, SEC Opens Informal Inquiry Of Whole Foods CEO Postings, Wall St. J., Jul. 14, 2007.
[48] Stephen Taub, Whole Foods Blogging Probe Dropped by SEC, CFO.com, Apr. 28, 2008.
[49] SEC v. Lidingo, Case 1:17-cv-02540, Apr. 10, 2017, https://www.sec.gov/litigation/complaints/2017/comp-pr2017-79-a.pdf

5.3 Intermediary Liability for Manipulative Attacks

While it might be difficult to hold pseudonymous attackers directly accountable, the same cannot be said for intermediaries like Seeking Alpha. Seeking Alpha not only hosts the content of pseudonymous attacks but also holds the keys to fictitious accounts. By requiring that these accounts be password-protected and linked to some external point of contact (e.g., e-mail address), Seeking Alpha keeps pseudonymous accounts from being hijacked by anonymous impersonators, claiming that they "insist on receiving the author's real name and contact information (which we keep confidential) and maintain a correspondence with the author."[50] Moreover, Seeking Alpha presumably records the IP address of the pseudonymous poster in its web logs, though that address may be unreliable if the author uses a service like Tor.

As gatekeepers of the link between pseudonymous accounts and underlying authors, intermediaries like Seeking Alpha are well-suited to punish systematic manipulators without chilling pseudonymity itself. One could easily imagine a policy whereby Seeking Alpha prohibits authors from opening new pseudonymous accounts after demonstrating a history of price reversals or options trading in excess of a given threshold. That sort of rule would promote the benefits of pseudonymity while holding authors accountable for price-distorting behavior, even when manipulative intent cannot be directly proven.

Could intermediaries like Seeking Alpha be induced to adopt this sort of policy by existing law? Much of the policy conversation around online intermediaries has focused on Section 230 of the Communications Decency Act, which provides that "[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider."[51] This statute immunizes intermediaries from defamation liability, and indeed one court has already rejected a defamation claim against Seeking Alpha on the basis of Section 230.[52] It is likely that Section 230 would also foreclose a direct claim of securities fraud or market manipulation against Seeking Alpha, because Seeking Alpha would not be considered the "speaker" of any misstatement.

However, it is less clear that Seeking Alpha would be immune from secondary liability for

[50] https://seekingalpha.com/page/policy_anonymous_contributors
[51] 47 U.S.C. 230.
[52] Nordlicht v. Seeking Alpha, Inc., et al., No. 64319/15 (N.Y. Sup. Ct., May 2, 2016).

aiding and abetting the manipulation documented here. In *SEC v. Apuzzo*, the Second Circuit delineated the elements of aiding and abetting liability: "(1) the existence of a securities law violation by the primary (as opposed to the aiding and abetting) party; (2) knowledge of this violation on the part of the aider and abettor; and (3) 'substantial assistance' by the aider and abettor in the achievement of the primary violation."[53] The knowledge element was relaxed by the Dodd-Frank Act to encompass recklessness.[54] An aiding-and-abetting claim may only be brought by the SEC, not a private plaintiff,[55] and in the Second Circuit, the SEC need not prove that the aider and abettor proximately caused the primary securities law violation.[56]

Suppose the SEC were able to establish existence of a securities law violation by a pseudonymous author — a nontrivial challenge to be sure, but this burden may be met in some cases. Does Seeking Alpha have knowledge of the manipulation, or at least a mens rea rising to the level of recklessness? It is unclear whether a court will hold Seeking Alpha to have knowledge of stock-price reversals like the ones documented in this article or even that the reversals "were so obvious that [Seeking Alpha] must have been aware."[57] On the one hand, this information is publicly available, and as this article has shown, it is not difficult to find anecdotal examples of reversals induced by pseudonymous attackers.

On the other hand, it is unclear that a court will effectively require Seeking Alpha to monitor stock prices for the hundreds if not thousands of publicly traded firms that are the subject of both positive and negative articles. It would seem that the technological barriers to this sort of monitoring are low, but legitimate questions might be raised concerning, for example, the degree of statistical confidence that is required to conclude that a given author is engaging in this sort of manipulation. Must Seeking Alpha demonstrate that the price reversal was unlikely to have been caused by random chance at, say, the 5% significance level? Courts might hesitate to rule in such a manner that would effectively impose this sort of affirmative

[53] 689 F.3d 204, 211 (2d. Cir. 2012).

[54] 15 U.S.C. 78t(e) ("For purposes of any action brought by the Commission under paragraph (1) or (3) of section 78u(d) of this title, any person that knowingly or recklessly provides substantial assistance to another person in violation of a provision of this chapter, or of any rule or regulation issued under this chapter, shall be deemed to be in violation of such provision to the same extent as the person to whom such assistance is provided.").

[55] 15 U.S.C. 78t(e); Central Bank v. First Interstate Bank, 511 U.S. 164 (1994).

[56] 689 F.3d at 213.

[57] SEC v. Wey, 246 F.Supp. 3d 894 (2017) (quoting Novak v. Kasaks, 216 F.3d 300, 308 (2d Cir. 2000).).

monitoring mandate absent SEC rulemaking or legislative action.

The final question is whether Seeking Alpha provided "substantial assistance" in the achievement of the primary violation. The Second Circuit has defined this term as requiring that the adier-and-abettor "in some sort associate[d] himself with the venture, that he participate[d] in it as in something that he wishe[d] to bring about, [and] that he [sought] by his action to make it succeed."[58] It is unclear whether merely publishing manipulative pseudonymous attacks constitutes the provision of substantial assistance. Note, however, that Seeking Alpha also maintains identification information and password protects these accounts, preventing anonymous impersonators from usurping these pseudonyms. This may very well constitute the provision of substantial assistance under *Apuzzo*.

5.4 SEC Rulemaking and Pseudonymous Attacks

There are two ways that SEC rulemaking could mitigate this sort of short-and-distort without chilling pseudonymous speech more generally. The first is for the SEC to promulgate a safe harbor for intermediaries like Seeking Alpha. This safe harbor would provide that an intermediary which bans authors who have repeatedly published attacks followed by reversals, and engaged in account-switching, would not be liable for the provision of substantial assistance to any underlying violation. Such a safe harbor would not prohibit identity-switching outright but encourage online intermediaries to discipline pseudonymous manipulation.

The second is for the SEC to impose an affirmative duty on online intermediaries like Seeking Alpha to maintain identifying information for pseudonymous accounts. This would ease the burden on enforcement authorities to prosecute these sort of manipulative attacks. Indeed, Seeking Alpha already claims that they "insist on receiving the author's real name and contact information" which is kept in confidence, and they "maintain a correspondence with the author, forwarding the author any questions or concerns that may emerge about their articles."[59] It is unclear how accurate these records are, so SEC rulemaking here would likely enhance the effectiveness of enforcement investigations against pseudonymous short sellers.

[58] *Id.* at 206 (quoting United States v. Peoni, 100 F.2d 401, 402 (2d Cir.1938)).
[59] https://seekingalpha.com/page/policy_anonymous_contributors

6 Conclusion

This paper shows that in financial markets, pseudonymity facilitates profitable manipulation of stock prices. Pseudonymous authors publish negative rumors about public companies that lead to significant short-term trading profits—and sharp reversals of the stock price decline. When markets realize that the pseudonymous author is spreading baseless rumors, the author switches to a new pseudonym, repeating the pattern. As gatekeepers of the link between pseudonymous accounts and underlying authors, online intermediaries like Seeking Alpha are well-suited to punish systematic manipulators without chilling pseudonymity itself. Regulators should consider a safe harbor for intermediaries who discipline repeat manipulators, as well as an affirmative duty to maintain identifying information for pseudonymous accounts.

References

An, Byeong-Je, Ang, Andrew, Bali, Turan G, & Cakici, Nusret. 2014. The joint cross section of stocks and options. *The Journal of Finance*, **69**(5), 2279–2337.

Beber, Alessandro, & Pagano, Marco. 2013. Short-selling bans around the world: Evidence from the 2007–09 crisis. *The Journal of Finance*, **68**(1), 343–381.

Benabou, Roland, & Laroque, Guy. 1992. Using privileged information to manipulate markets: Insiders, gurus, and credibility. *The Quarterly Journal of Economics*, **107**(3), 921–958.

Boehmer, Ekkehart, Jones, Charles M, & Zhang, Xiaoyan. 2013. Shackling short sellers: The 2008 shorting ban. *The Review of Financial Studies*, **26**(6), 1363–1400.

Campbell, John L, DeAngelis, Matthew D, & Moon, James. 2017. The Credibility of the Crowds: Personal Stock Holdings and Investors Response to Stock Opinions on Social Media.

Cao, Charles, Chen, Zhiwu, & Griffin, John M. 2005. Informational content of option volume prior to takeovers. *The Journal of Business*, **78**(3), 1073–1109.

Chakravarty, Sugato, Gulen, Huseyin, & Mayhew, Stewart. 2004. Informed trading in stock and option markets. *The Journal of Finance*, **59**(3), 1235–1257.

Comerton-Forde, Carole, & Putniņš, Tālis J. 2008. Are short sellers manipulating the market? *Wall Street Journal*.

Cremers, Martijn, & Weinbaum, David. 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, **45**(2), 335–367.

Easley, David, O'hara, Maureen, & Srinivas, Pulle Subrahmanya. 1998. Option volume and stock prices: Evidence on where informed traders trade. *The Journal of Finance*, **53**(2), 431–465.

Farrell, Joseph, & Rabin, Matthew. 1996. Cheap talk. *Journal of Economic perspectives*, **10**(3), 103–118.

Fox, Merritt B, Glosten, Lawrence R, & Rauterberg, Gabriel V. 2018. Stock Market Manipulation and Its Regulation. *Yale J. on Reg.*, **35**, 67.

Glosten, Lawrence R, & Milgrom, Paul R. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, **14**(1), 71–100.

Glosten, Lawrence R, & Putniņš, Tālis J. 2016. Welfare Costs of Informed Trade.

Grimmer, Justin, & Stewart, Brandon M. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, **21**(3), 267–297.

Grossman, Sanford J. 1981. The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics*, **24**(3), 461–483.

Holmes, David I. 1992. A stylometric analysis of Mormon scripture and related texts. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 91–120.

Iqbal, Farkhund, Binsalleeh, Hamad, Fung, Benjamin CM, & Debbabi, Mourad. 2010. Mining writeprints from anonymous e-mails for forensic investigation. *digital investigation*, **7**(1-2), 56–64.

Jayaraman, Narayanan, Frye, Melissa B, & Sabherwal, Sanjiv. 2001. Informed trading around merger announcements: An empirical test using transaction volume and open interest in options market. *Financial Review*, **36**(2), 45–74.

Kogan, Shimon, Moskowitz, Tobias J, & Niessner, Marina. 2018. Fake News in Financial Markets.

Macey, Jonathan, & Mitts, Joshua. 2014. Finding order in the morass: The three real justifications for piercing the corporate veil. *Cornell L. Rev.*, **100**, 99.

Meulbroek, Lisa K. 1992. An empirical analysis of illegal insider trading. *The Journal of Finance*, **47**(5), 1661–1699.

Mitts, Joshua, & Talley, Eric L. 2018. Informed Trading and Cybersecurity Breaches.

Mosteller, Frederick, & Wallace, David. 1964. Inference and disputed authorship: The Federalist.

Narayanan, Arvind, Paskov, Hristo, Gong, Neil Zhenqiang, Bethencourt, John, Stefanov, Emil, Shin, Eui Chul Richard, & Song, Dawn. 2012. On the feasibility of internet-scale author identification. *Pages 300–314 of: Security and Privacy (SP), 2012 IEEE Symposium on*. IEEE.

Pan, Jun, & Poteshman, Allen M. 2006. The information in option volume for future stock prices. *The Review of Financial Studies*, **19**(3), 871–908.

Publius. 1787. The Federalist. *The Independent Journal, New York Packet, Daily Advertiser.*

Putniņš, Tālis J. 2012. Market manipulation: A survey. *Journal of Economic Surveys*, **26**(5), 952–967.

Renaulta, Thomas. 2018. Pump-and-dump or news? Stock market manipulation on social media.

Saffi, Pedro AC, & Sigurdsson, Kari. 2010. Price efficiency and short selling. *The Review of Financial Studies*, **24**(3), 821–852.

Tetlock, Paul C. 2011. All the news that's fit to reprint: Do investors react to stale information? *The Review of Financial Studies*, **24**(5), 1481–1512.

Van Bommel, Jos. 2003. Rumors. *The Journal of Finance*, **58**(4), 1499–1520.

Varian, Hal R. 2014. Big data: New tricks for econometrics. *Journal of Economic Perspectives*, **28**(2), 3–28.

Vila, Jean-Luc. 1989. Simple games of market manipulation. *Economics Letters*, **29**(1), 21–26.

Wong, Yu Ting Forester, & Zhao, Wuyang. 2017. Post-Apocalyptic: The Real Consequences of Activist Short-Selling.

Zhao, Wuyang. 2018. Activist short-selling.

Zou, Hui, & Hastie, Trevor. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **67**(2), 301–320.

7 Tables

Table 1: Pseudonymous and Non-Pseudonymous Authors

This table presents selected examples of pseudonymous and non-pseudonymous authors in my estimation dataset.

Pseudonymous Authors	Non-Pseudonymous Authors
Midnight Trader	Kevin Quon
Bargain Bin	Cliff Wachtel
Alpha Generator	Citron Research
Follow The Data	Josh Young
Tweakerlabs	Gary Weiss
AlchemyOfFinance	Akshay Kaul
Disruptive Investor	Philip Davis
Vatalyst	David Urban
Efficient Alpha	Larry MacDonald
BumbleBayGoombeeFluor	Joseph Bohm

Table 2: Summary Statistics

This table reports summary statistics for the continuous variables in the primary sample of 4,785 firm-articles by firms with \$2 million in market capitalization or more. Categorical variables like firm industry are not included in this table, but are included as described elsewhere in the paper.

Variable	N	Mean	Std. Dev.	Min	Max	25%	50%	75%
car_{i,j,t_0+2,t_0+5}	4,784	.001	.048	-.722	.4	-.018	.001	.021
car_{i,t_0-1,t_0+1}	4,785	-.006	.069	-1.94	1.229	-.022	-.001	.017
$rev_{i,j,t}$	4,784	.007	.083	-1.107	1.923	-.025	.004	.033
$rev_{i,j,t} > 0$	4,785	.535	.499	0	1	0	1	1
$rev_{i,j,t} > 0.02$	4,785	.346	.476	0	1	0	0	1
$\Delta spread_{i,t_0,t_0+2}$	4,772	.335	2.313	-1	80.919	0	0	0
$\Delta sp\hat{r}ead_{i,t_0,t_0+2}$	4,772	.336	2.31	-1	79.229	-.034	-.001	.031
Market Value (in \$1 mil.)	4,785	58,518	104,994	2,013	626,550	4,702	14,078	55,930
Total Assets	4,785	96,405	329,839	1.001	2,807,491	3,091	12,548	51,839
Total Liabilities	4,736	77,085	301,654	.128	2,736,580	1,807	7,584	33,269
Net Income	4,785	3,183	7,685	-14,685	53,394	23.767	373	2,856
Amihud (2002) Illiquidity	4,785	.001	.018	0	1.139	0	0	0
Idiosyncratic Volatility	4,785	.024	.017	.005	.309	.014	.02	.028
Article Hour	4,785	10.502	4.75	0	23	7	10	14
Year	4,785	2014	2001	2010	2017	2012	2014	2016

Table 3: Predictors of Pseudonymous Authorship: Full (Unmatched) Sample

This table examines univariate predictors of pseudonymous authorship on firm- and article-level covariates prior to employing propensity-score matching. For each covariate, the table reports the mean for the pseudonymous articles, the mean for non-pseudonymous articles, the difference in means as a percentage of the non-pseudonymous mean, the t-statistic for that difference, and the p-value of that t-statistic.

Variable	Pseudonymous	Real Name	% bias	t-statistic	$p > \|t\|$
Market Value (in $1 mil.)	51,195	62,692	-11.3	-3.65	0
Total Assets	93,556	98,028	-1.4	-0.45	0.652
Total Liabilities	75,682	77,885	-0.7	-0.24	0.809
Net Income	2,740	3,436.1	-9.3	-3.02	0.003
Amihud (2002) Illiquidity	0.00093	0.00037	2.7	1.04	0.296
Idiosyncratic Volatility	0.02404	0.02344	3.5	1.15	0.25
Article Hour	10.503	10.501	0	0.01	0.993
Year	2,013.9	2,013.9	-3.7	-1.24	0.214
Industry: Materials	0.04147	0.03483	3.5	1.17	0.244
Industry: Capital Goods	0.04896	0.03911	4.8	1.62	0.105
Industry: Commercial & Professional Services	0.00403	0.0069	-3.9	-1.25	0.211
Industry: Transportation	0.02477	0.01906	3.9	1.32	0.187
Industry: Automobiles & Components	0.02362	0.03385	-6.1	-1.99	0.047
Industry: Consumer Durables & Apparel	0.04205	0.03089	6	2.02	0.043
Industry: Consumer Services	0.07604	0.06901	2.7	0.91	0.365
Industry: Media	0.01325	0.0253	-8.8	-2.8	0.005
Industry: Retailing	0.10253	0.14558	-13.1	-4.26	0
Industry: Food & Staples Retailing	0.01037	0.01249	-2	-0.65	0.513
Industry: Food, Beverage & Tobacco	0.04032	0.03943	0.5	0.15	0.88
Industry: Household & Personal Products	0.01671	0.02596	-6.4	-2.07	0.038
Industry: Health Care Equipment & Services	0.02247	0.01249	7.6	2.64	0.008
Industry: Pharma, Biotech & Life Sciences	0.07028	0.03023	18.4	6.46	0
Industry: Banks	0.02995	0.03319	-1.9	-0.61	0.541
Industry: Diversified Financials	0.0265	0.02202	2.9	0.98	0.327
Industry: Insurance	0.00634	0.00296	5	1.74	0.082
Industry: Software & Services	0.1394	0.1791	-10.9	-3.56	0
Industry: Technology Hardware & Equipment	0.06509	0.07756	-4.8	-1.59	0.111
Industry: Semiconductors & Equipment	0.06336	0.05587	3.2	1.06	0.289
Industry: Telecommunication Services	0.02247	0.03418	-7.1	-2.29	0.022
Industry: Utilities	0.01325	0.00526	8.4	2.96	0.003
Industry: Real Estate	0.01152	0.00789	3.7	1.26	0.206

Table 4: Matched Sample: Balance Test

This table examines whether the matching yields a balanced sample for firm- and article-level covariates between the pseudonymous and non-pseudonymous articles. For each covariate, the table reports the mean for the pseudonymous articles, the mean for non-pseudonymous articles, the difference in means as a percentage of the non-pseudonymous mean, the t-statistic for that difference, and the p-value of that t-statistic. As the table shows, none of the p-values are below the 5% significance level, indicating the differences in means are not statistically significant.

Variable	Pseudonymous	Real Name	% bias	t-statistic	$p > \|t\|$
Market Value (in $1 mil.)	51,545	51,576	0	-0.01	0.992
Total Assets	94,452	85,140	2.8	0.89	0.371
Total Liabilities	75,726	66,491	3.1	0.98	0.329
Net Income	2,765.1	2,891.1	-1.7	-0.53	0.595
Amihud (2002) Illiquidity	0.00094	0.00026	3.3	0.98	0.326
Idiosyncratic Volatility	0.02408	0.02476	-4	-1.06	0.291
Article Hour	10.515	10.49	0.5	0.15	0.878
Year	2013.9	2013.8	3.9	1.13	0.258
Fraud-Related Text (Word & Phrase Count)	.2936	.33953	-2.3	-0.82	0.412
Industry: Materials	0.04186	0.04942	-3.9	-1.06	0.288
Industry: Capital Goods	0.04942	0.03779	5.6	1.67	0.095
Industry: Commercial & Professional Services	0.00407	0.0064	-3.1	-0.95	0.345
Industry: Transportation	0.025	0.03663	-7.9	-1.97	0.048
Industry: Automobiles & Components	0.02384	0.02616	-1.4	-0.44	0.662
Industry: Consumer Durables & Apparel	0.04244	0.04477	-1.2	-0.33	0.738
Industry: Consumer Services	0.06919	0.07151	-0.9	-0.27	0.79
Industry: Media	0.01337	0.01337	0	0	1
Industry: Retailing	0.10349	0.10465	-0.4	-0.11	0.911
Industry: Food & Staples Retailing	0.01047	0.00988	0.5	0.17	0.865
Industry: Food, Beverage & Tobacco	0.0407	0.0407	0	0	1
Industry: Household & Personal Products	0.01686	0.01163	3.6	1.29	0.195
Industry: Health Care Equipment & Services	0.02267	0.01977	2.2	0.59	0.554
Industry: Pharma, Biotech & Life Sciences	0.07093	0.07151	-0.3	-0.07	0.947
Industry: Banks	0.03023	0.02442	3.3	1.05	0.296
Industry: Diversified Financials	0.02674	0.03023	-2.3	-0.61	0.539
Industry: Insurance	0.0064	0.01047	-6	-1.31	0.192
Industry: Software & Services	0.13895	0.13721	0.5	0.15	0.882
Industry: Technology Hardware & Equipment	0.0657	0.06686	-0.4	-0.14	0.891
Industry: Semiconductors & Equipment	0.06395	0.06628	-1	-0.28	0.782
Industry: Telecommunication Services	0.02267	0.02093	1	0.35	0.726
Industry: Utilities	0.01337	0.0093	4.2	1.13	0.26
Industry: Real Estate	0.01163	0.01105	0.6	0.16	0.872

Table 5: Stock Price Reversals: Pseudonymous vs. Non-Pseudonymous Articles

This table examines whether pseudonymous articles are followed by greater stock-price reversals than non-pseudonymous articles by examining four different specifications. Column (1) follows Tetlock (2011) and considers whether the correlation between car_{i,j,t_0+2,t_0+5}, i.e., the four-factor cumulative abnormal return (CAR) over the interval $[t-2, t+5]$, and car_{i,t_0-1,t_0+1}, the four-factor CAR over the interval $[t_0-1, t_0+1]$, differs between pseudonymous and non-pseudonymous articles. Columns (2)-(4) examine $rev_{i,j,t} = car_{i,j,t_0+2,t_0+5} - car_{i,j,t_0-1,t_0+1}$. Column (2) regresses $rev_{i,j,t}$ on an indicator for pseudonymous articles, column (3) regresses an indicator equal to 1 if $rev_{i,j,t} > 0$, and column (4) examines the cases where where $rev_{i,j,t} > 0.02$. All regressions employ propensity-score matching with treatment-control pairs as OLS regression weights and robust standard errors.

	car_{i,j,t_0+2,t_0+5}	$rev_{i,j,t}$	$rev_{i,j,t} > 0$	$rev_{i,j,t} > 0.02$
Pseudonymous $\times$ car_{i,t_0-1,t_0+1}	**-0.1139****			
	(-2.43)			
Pseudonymous Author	**0.0055*****	**0.0080****	**0.0465****	**0.0430****
	(2.73)	**(2.51)**	**(2.28)**	**(2.23)**
car_{i,t_0-1,t_0+1}	0.1024***			
	(2.87)			
(Intercept)	-0.0023	0.0030	0.5047***	0.3262***
	(-1.45)	(1.18)	(30.50)	(21.14)
Observations	2,899	2,899	2,899	2,899

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Pseudonymous Attacks and Options Trading

This table examines whether pseudonymous articles are followed by greater call options trading than put options trading on the day of publication t_0 compared to $[t_0 - 9, t_0 - 1]$, and during the period of the stock-price correction, that is, $[t_0{+}2, t_0{+}5]$ compared to $[t_0, t_0{+}1]$. I estimate a difference-in-difference-in-differences model, which compares the difference in log open interest or log volume between put and call options (difference #1) in the over-time change (difference #2) between pseudonymous and non-pseudonymous articles (difference #3). The dataset consists of nearly at-the-money options written on firms in the matched sample with an absolute delta between 0.45 and 0.55. The estimation employs firm-article fixed effects, and robust standard errors are clustered by firm-article.

	Open Interest		Volume	
	(1)	(2)	(3)	(4)
Pseudonymous × Publication Day × Call Option	**-0.0766****		**-0.0775****	
	(-2.41)		**(-2.23)**	
Pseudonymous × Correction Period × Call Option		**0.0892*****		**0.0620****
		(3.49)		**(2.04)**
Pseudonymous × Publication Day	0.0489*		-0.0006	
	(1.90)		(-0.02)	
Pseudonymous × Call Option	-0.0210	-0.0446	0.0130	-0.0277
	(-0.64)	(-1.10)	(0.54)	(-0.86)
Publication Day × Call Option	0.0251		0.0548**	
	(1.02)		(1.97)	
Pseudonymous × Correction Period		-0.0560***		-0.0077
		(-2.80)		(-0.25)
Correction Period × Call Option		-0.0316		-0.0171
		(-1.59)		(-0.72)
Publication Day	-0.0254		0.0834***	
	(-1.22)		(2.72)	
Correction Period (in $[t-2, t+5]$)		0.0451***		-0.0887***
		(2.85)		(-3.77)
Call Option	0.4954***	0.4918***	0.4763***	0.5026***
	(19.97)	(15.24)	(24.89)	(19.49)
| Delta |	-2.1201***	-2.4260***	-2.4397***	-2.4949***
	(-23.42)	(-21.50)	(-26.71)	(-21.17)
Time-to-Expiration	-0.0006***	-0.0006***	-0.0031***	-0.0031***
	(-8.03)	(-6.92)	(-37.97)	(-34.20)
Log of Strike Price	-1.0745***	-1.3208***	-1.3427***	-1.4128***
	(-5.72)	(-4.69)	(-8.51)	(-5.33)
Gamma	-1.1147***	-1.2278***	2.8839***	2.9693***
	(-9.88)	(-10.50)	(15.57)	(16.90)
(Intercept)	18.3620***	21.2865***	19.9068***	20.7935***
	(8.72)	(6.74)	(11.21)	(6.96)
Observations	578,381	313,695	403,609	219,923

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Pseudonymous Attacks and Options Characteristics

This table examines whether the abnormal options trading on the day of publication of pseudonymous articles and during the period of the stock-price correction is concentrated in options with unusual characteristics. I estimate the same model as Table 6, considering the outcomes of time-to-expiration, strike price, absolute delta and gamma. The estimations employ firm-article fixed effects, and robust standard errors are clustered by firm-article.

Panel A: Publication Day

	TTE	Strike Price	| Delta |	Gamma
Pseudonymous × Publication Day × Call Option	**-0.0470**	**-0.0008**	**0.0001**	**0.0007***
	(-0.05)	**(-1.16)**	**(0.11)**	**(1.66)**
Pseudonymous × Publication Day	-2.3974*	0.0072*	-0.0056**	0.0001
	(-1.71)	(1.91)	(-1.97)	(0.51)
Pseudonymous × Call Option	0.1597	0.0023**	-0.0000	-0.0005
	(0.26)	(2.40)	(-0.24)	(-1.64)
Publication Day × Call Option	0.1860	0.0006	-0.0001	-0.0004
	(0.23)	(1.20)	(-0.20)	(-1.41)
Publication Day	1.6128	-0.0061*	-0.0000	0.0004
	(1.32)	(-1.86)	(-0.20)	(0.64)
Call Option	-0.0207	-0.0061***	0.0003*	0.0023***
	(-0.04)	(-7.62)	(1.89)	(10.06)
(Intercept)	186.8216***	11.2191***	0.4999***	0.0599***
	(1192.03)	(39049.44)	(10458.51)	(715.55)
Observations	686,809	686,809	686,809	686,809

Panel B: Correction Period

	TTE	Strike Price	| Delta |	Gamma
Pseudonymous × Publication Day × Call Option	**-0.7423**	**0.0003**	**0.0001**	**0.0001**
	(-0.85)	**(0.55)**	**(0.20)**	**(0.43)**
Pseudonymous × Correction Period	2.2532*	0.0058***	-0.0000	-0.0003
	(1.83)	(2.77)	(-0.07)	(-0.38)
Pseudonymous × Call Option	0.8266	0.0019*	0.0001	-0.0002
	(0.94)	(1.66)	(0.21)	(-0.61)
Correction Period × Call Option	0.4804	-0.0001	-0.0004	-0.0000
	(0.72)	(-0.18)	(-1.09)	(-0.08)
Correction Period	-0.7897	-0.0022	0.0002	-0.0001
	(-0.82)	(-1.49)	(1.15)	(-0.17)
Call Option	-0.4832	-0.0060***	0.0004	0.0022***
	(-0.69)	(-6.73)	(1.46)	(7.31)
(Intercept)	189.0077***	11.2178***	0.4997***	0.0612***
	(445.73)	(15235.01)	(5333.84)	(239.18)
Observations	369,825	369,825	369,825	369,825

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Pseudonymous Attacks and Put-Call Parity

This table examines whether pseudonymous articles are followed by greater deviations from put-call parity over the window $[t_0, t_0+2]$ compared to $[t_0-9, t_0-1]$, and during the period following the acquisition of call options, that is, $[t_0+3, t_0+5]$ compared to $[t_0, t_0+2]$. The estimations are lagged by one day in contrast to Table 6, in order to account for options markets updating in response to informed order flow. In the spirit of Cremers & Weinbaum (2010), I employ fixed effects for put and call options written on the same underlying security, expiration date and strike price. I estimate a difference-in-difference-in-differences model, which compares the difference in implied volatility between put and call options (difference #1) in the over-time change (difference #2) between pseudonymous and non-pseudonymous articles (difference #3). The dataset consists of nearly at-the-money options written on firms in the matched sample with an absolute delta between 0.45 and 0.55. Robust standard errors are clustered by the unique combination of security-expiration date-strike price.

	(1)	(2)	(3)	(4)
Pseudonymous × $[t_0, t_0+2]$ × **Call Option**	**-0.0018*** **	**-0.0017*** **		
	(-3.68)	**(-3.48)**		
Pseudonymous × $[t_0+3, t_0+5]$ × **Call Option**			**0.0023*** **	**0.0022*** **
			(4.60)	**(4.67)**
Pseudonymous × $[t_0, t_0+2]$	0.0028***	0.0023***		
	(3.94)	(3.42)		
Pseudonymous × Call Option	0.0047***	0.0045***	0.0030***	0.0029***
	(11.68)	(11.40)	(5.41)	(5.45)
$[t_0, t_0+2]$ × Call Option	0.0007**	0.0006*		
	(1.96)	(1.73)		
Pseudonymous × $[t_0+3, t_0+5]$			0.0002	-0.0000
			(0.30)	(-0.04)
$[t_0+3, t_0+5]$ × Call Option			-0.0007*	-0.0006*
			(-1.74)	(-1.68)
$[t_0, t_0+2]$	0.0003	0.0007		
	(0.50)	(1.52)		
$[t_0+3, t_0+5]$			-0.0024***	-0.0010**
			(-5.07)	(-2.19)
Call Option	-0.0129***	-0.0125***	-0.0116***	-0.0112***
	(-33.39)	(-33.27)	(-23.96)	(-23.73)
| Delta |		0.0020		-0.0021
		(0.88)		(-0.70)
Gamma		-0.3565***		-0.4393***
		(-14.31)		(-9.90)
Vega		0.0003***		0.0003***
		(7.99)		(5.78)
Theta		-0.0005***		-0.0004***
		(-14.93)		(-9.42)
(Intercept)	0.4238***	0.4205***	0.4171***	0.4248***
	(633.47)	(146.05)	(438.11)	(104.66)
Observations	813,599	813,599	369,825	369,825

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Words and Phrases Predicting Pseudonymous Attacks

This table presents the top 100 stemmed words and phrases which are most predictive of pseudonymous authorship, as estimated by a logistic regression with elastic-net regularization (Zou & Hastie, 2005). The elastic-net parameter λ is chosen by ten-fold cross-validation, maximizing the area under the ROC curve. The in-sample accuracy of the model is 0.9000, with a sensitivity (true positive rate) of 0.7568 and a specificity (true negative rate) of 0.9983. The words and phrases below seem to reflect idiosyncratic noise, suggesting that pseudonymous authors are manipulating markets by engaging in informed options trading, not by posting provocative content.

Term	Coefficient	Term	Coefficient	Term	Coefficient	Term	Coefficient
invest longshort	1.51	save stock	0.27	name stock	0.16	upsid potenti	0.09
short onli special	1.35	individu stock	0.27	financi media	0.16	tailor	0.09
invest follow	1.33	traderstop	0.26	secur mention thi	0.16	price valu	0.09
technician	0.73	equiti portfolio	0.24	competitor alreadi	0.16	current hold	0.09
thi articl wrien	0.70	thing chang	0.23	strategi data	0.16	segment also	0.08
firepow analyt	0.69	stock befor	0.23	hedg posit	0.15	hub stock	0.08
research longshort	0.65	investor long onli	0.22	alertstop	0.15	author stock	0.08
futur articl	0.64	represent made	0.22	larg corpor	0.15	world will	0.08
secur mention	0.55	razor	0.22	let take closer	0.14	luckili	0.07
onli special	0.50	pricelin nasdaqpcln	0.21	increas invest	0.14	august	0.07
valu valu special	0.47	disclosur amw short	0.21	higher qualiti	0.14	weigh machin	0.07
follow followingy	0.39	investor view	0.20	reason price longterm	0.13	sell rate	0.07
monei dure	0.38	comment short	0.20	demand product	0.13	ani secur mention	0.07
invest summari	0.37	research invest	0.20	asset return asset	0.11	qineqt	0.06
financi properti casualti	0.35	thi alon	0.20	believ look	0.11	captur oct	0.06
trader summari	0.33	bank unit state	0.19	next six	0.11	conflict interest	0.06
invest tech	0.33	valuat current	0.19	kraken	0.11	idea fund	0.06
investor summari	0.31	recap	0.19	pai premium	0.11	share yield	0.06
start point	0.30	technolog	0.19	consensu estim	0.10	financi servic	0.06
now good	0.30	trade rang	0.18	bid etc	0.10	doctorx	0.06
global economi	0.30	idea longshort	0.18	contact full	0.10	invest advisor thi	0.06
messagescout	0.29	follow two	0.18	thi articl inform	0.10	certain point	0.06
analyst longshort	0.28	valu investor	0.17	bearish view	0.10	takeov analyst	0.06
earn degre	0.28	abov fair	0.17	dividend yield	0.10	uncertain	0.06
market see	0.27	articl wrien	0.17	messag market	0.09	exponenti	0.05

Table 10: Pseudonymous Attacks and Bid/Ask Spreads

This table examines how bid/ask spreads respond to the publication of pseudonymous attack articles which are expected to be followed by informed options trading and price reversals. Columns (1)-(3) examine the percentage change in the raw bid-ask spread from t_0 to $t+2$, and columns (4)-(6) normalize the spread by the price of the underlying stock on day t_0 and $t+2$. All models report the results of regressing the change in the spread on an indicator equal to 1 if the article was written by a pseudonymous article, interacted with the cumulative abnormal log return over the period $[t_0 - 1, t_0 + 1]$. All regressions employ propensity-score matching with treatment-control pairs as OLS regression weights with standard errors robust to heteroskedasticity.

	$\Delta spread_{i,t,\tau}$			$\Delta \widehat{spread}_{i,t,\tau}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Pseudonymous $\times$ car_{i,t_0-1,t_0+1}	**-2.1793***			**-2.2366***		
	(2.89)			**(2.92)**		
Pseudonymous $\times$ $car_{i,t_0-1,t_0+1} < 0$		**0.4352***			**0.4479***	
		(2.48)			**(2.55)**	
Pseudonymous $\times$ $car_{i,t_0-1,t_0+1} < -0.05$			**0.5037**			**0.5180**
			(2.46)			**(2.39)**
Pseudonymous Author	-0.0789	-0.1821	-0.1242	-0.0824	-0.1877*	-0.1286
	(-0.91)	(-1.62)	(-1.31)	(-0.96)	(-1.69)	(-1.38)
car_{i,t_0-1,t_0+1}	1.3319***			1.1158**		
	(2.65)			(2.22)		
$car_{i,t_0-1,t_0+1} < 0$		-0.2897**			-0.2626**	
		(-2.56)			(-2.33)	
$car_{i,t_0-1,t_0+1} < -0.05$			-0.2210*			-0.1828
			(-1.96)			(-1.62)
(Intercept)	0.3354***	0.4074***	0.3539***	0.3373***	0.4031***	0.3525***
	(4.29)	(3.89)	(4.08)	(4.36)	(3.89)	(4.11)
Observations	2,893	2,893	2,893	2,893	2,893	2,893

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Pseudonymous Attacks and Net Trading Losses to Sellers

This table calculates the aggregate trading losses to sellers as a result of the depressed prices induced by pseudonymous attacks. For each firm-article written by pseudonymous authors, I calculate the total dollar volume on each of the trading days in the interval $[t_0, t+4]$ and compute a counterfactual dollar volume consisting of the number of trades multiplied by the price at $t_0 + 5$. This is the price that sellers would have received in the absence of any price distortion, i.e., if the shares had been sold at their price on day $t_0 + 5$. I subtract the actual dollar volume from the counterfactual to derive the net losses to sellers. These losses have zero welfare effects ex post: they consist solely of transfers, and thus the gains to buyers exactly equal the losses to sellers.

	t_0	$t+1$	$t+2$	$t+3$	$t+4$
Actual Dollar Volume	$1.044 trillion	$976 billion	$977 billion	$950 billion	$933 billion
Counterfactual Dollar Volume	$1.050 trillion	$981 billion	$982 billion	$954 billion	$935 billion
Net Losses to Sellers (Counterfactual - Actual)	$5.40 billion	$4.62 billion	$4.22 billion	$3.96 billion	$1.93 billion
Grand Total of Net Losses to Sellers	**$20.1 billion**				

Table 12: Stock Price Reversals: Pseudonymous Authors in Trustworthy Periods

This table examines whether pseudonymous articles are followed by greater stock-price reversals than real-name articles by examining four different specifications in trustworthy vs. non-trustworthy periods (see text for definition). Columns (1) and (5) follow Tetlock (2011) and considers whether the correlation between car_{i,j,t_0+2,t_0+5}, i.e., the four-factor cumulative abnormal return (CAR) over the interval $[t-2, t+5]$, and car_{i,t_0-1,t_0+1}, the four-factor CAR over the interval $[t_0-1, t_0+1]$, differs between pseudonymous and non-pseudonymous articles. The other columns examine $rev_{i,j,t} = car_{i,j,t_0+2,t_0+5} - car_{i,t_0-1,t_0+1}$. Columns (2) and (6) regresses $rev_{i,j,t}$ on an indicator for pseudonymous articles, columns (3) and (7) regress an indicator equal to 1 if $rev_{i,j,t} > 0$, and columns (4) and (8) examine cases where $rev_{i,j,t} > 0.02$. All regressions employ robust standard errors.

	Trustworthy				Non-Trustworthy			
	car_{i,j,t_0+2,t_0+5}	$rev_{i,j,t}$	$rev_{i,j,t} > 0$	$rev_{i,j,t} > 0.02$	car_{i,j,t_0+2,t_0+5}	$rev_{i,j,t}$	$rev_{i,j,t} > 0$	$rev_{i,j,t} > 0.02$
Pseudonymous	**0.0047****	**0.0077****	**0.0567****	**0.0483****	**0.0025**	**0.0038**	**-0.0118**	**-0.0006**
Author	**(2.00)**	**(2.07)**	**(2.38)**	**(2.02)**	**(0.62)**	**(0.75)**	**(-0.34)**	**(-0.02)**
Pseudonymous	**-0.1351****				**0.0252**			
$\times\ car_{i,t_0-1,t_0+1}$	**(-2.40)**				**(0.26)**			
car_{i,t_0-1,t_0+1}	0.1426***				0.0782*			
	(3.17)				(1.83)			
(Intercept)	0.0032*	0.0173***	0.5887***	0.4080***	-0.0137***	-0.0323***	0.2978***	0.1247***
	(1.73)	(5.80)	(30.03)	(21.10)	(-4.62)	(-8.05)	(11.10)	(6.61)
Observations	2,107	2,107	2,107	2,107	792	792	793	793

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Pseudonymous Attacks and Options Trading: Trustworthy Periods

This table examines whether pseudonymous articles are followed by greater call options trading than put options trading on the day of publication t_0 and during the period of the stock-price correction, i.e., $[t_0 - 2, t_0 + 5]$, comparing between trustworthy and non-trustworthy periods (see text for definition). In every column, the outcome is the log open interest for a nearly at-the-money option with an absolute delta between 0.45 and 0.55. I estimate a difference-in-difference-in-differences model, which compares the difference in log open interest between put and call options (difference #1) in the over-time change (difference #2) between pseudonymous and non-pseudonymous articles (difference #3). The estimations employ firm-article fixed effects, and robust standard errors are clustered by firm-article.

	Trustworthy		Non-Trustworthy	
	(1)	(2)	(3)	(4)
Pseudonymous × Publication Day × Call Option	**-0.0803****		**-0.0399**	
	(-2.18)		**(-0.65)**	
Pseudonymous × Correction Period × Call Option		**0.0972*****		**0.0517**
		(3.17)		**(1.13)**
Pseudonymous × Publication Day	0.0365		0.0675	
	(1.23)		(1.33)	
Pseudonymous × Call Option	-0.0403	-0.0549	0.0422	0.0240
	(-1.03)	(-1.14)	(0.69)	(0.33)
Publication Day × Call Option	-0.0145		0.1155**	
	(-0.51)		(2.46)	
Pseudonymous × Correction Period		-0.0344		-0.1011***
		(-1.43)		(-2.93)
Correction Period × Call Option		-0.0019		-0.1043***
		(-0.08)		(-3.03)
Publication Day	0.0066		-0.1059***	
	(0.27)		(-2.61)	
Correction Period (in $[t-2, t+5]$)		0.0100		0.1318***
		(0.52)		(5.11)
Call Option	0.4780***	0.4160***	0.5355***	0.6684***
	(15.90)	(10.87)	(12.33)	(11.91)
(Intercept)	18.9861***	22.1288***	16.7316***	19.2136***
	(7.95)	(6.04)	(3.97)	(3.00)
Observations	418,834	226,764	159,547	86,931

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Pseudonymous Attacks and Disappearing Authors

This table examines three propositions discussed in the text. Column (1) examines whether pseudonymous authors are more likely to "disappear," i.e., whether a given article is likely to be the last one written by an author. Column (2) tests whether the market response to an article is linked to the presence of prior reversals. In this column, there is no observation for the first article by an author. Columns (3) and (4) link the two prior tests together and consider whether pseudonymous authors are more likely to "disappear" when the market has ceased to respond to the publication of an article. The cumulative abnormal return car_{i,j,t_0-1,t_0+1} is standardized. All regressions are estimated with robust standard errors on the matched sample with treatment-control pairs using OLS regression weights.

	$last_{i,j,t}$	$\lvert car_{i,j,t_0-1,t_0+1}\rvert$	$last_{i,j,t}$	$last_{i,j,t}$
Pseudonymous Author	**0.0384****		**0.0244**	
	(2.36)		**(1.41)**	
Prior Negative Reversals ($prior_{i,j,t}$)		**-0.2488*****		
		(-3.98)		
Pseudonymous Author ×			**-0.0794****	
$\lvert car_{i,j,t_0-1,t_0+1}\rvert$			**(-2.39)**	
$\lvert car_{i,j,t_0-1,t_0+1}\rvert$			0.0363	
			(1.28)	
Low-Credibility Pseudonymous				**0.7595*****
				(27.87)
(Intercept)	0.1895***	0.0336***	0.1954***	0.1944***
	(14.93)	(30.60)	(14.10)	(24.09)
Observations	2,900	2,601	2,900	2,900

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Stylometric Analysis of Pseudonymous Authorship

This table examines whether pseudonymous attacks are stylometrically more similar to articles written by "former authors." The term "former authors" is defined as authors who had written their last article before the publication date of the attack, exploiting the benefit of hindsight to determine this date. Column (1) examines this link unconditionally. Column (2) examines whether pseudonymous authors hide their identities when first appearing, but eventually revert to their true writing style, i.e., whether similarity to former authors increases as an author writes more articles. Columns (3) and (4) consider discrete versions of the continuous interaction in Column (2) at the 5th and 10th article, respectively. All regressions are estimated with robust standard errors on the matched sample with treatment-control pairs using OLS regression weights and include year fixed effects. The number of observations is lower than prior tables because the full text of some articles was unavailable on the Internet Archive.

	(1)	(2)	(3)	(4)
Pseudonymous Author	**0.0065***	**-0.0026**	**0.0185*****	**0.0213*****
	(1.67)	**(-0.56)**	**(3.70)**	**(3.73)**
Pseudonymous × **Article #**		**0.0006*****		
		(5.74)		
Pseudonymous × **Article #** < 5			**-0.0231*****	
			(-2.92)	
Pseudonymous × **Article #** < 10				**-0.0210*****
				(-2.70)
Article #		0.0000		
		(0.47)		
Article # < 5			-0.0012	
			(-0.18)	
Article # < 10				-0.0051
				(-0.81)
(Intercept)	0.0588***	0.0598***	0.0595***	0.0619***
	(7.23)	(7.25)	(7.03)	(6.80)
Observations	2,518	2,518	2,518	2,518

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

8 Online Appendix

Figure 7: Density of Propensity Score: Pseudonymous vs. Non-Pseudonymous

This figure shows the density of the propensity score for articles written by pseudonymous and non-pseudonymous authors. As the figure shows, the two groups are almost exactly balanced on the propensity score, consistent with the variable-by-variable balance tests in Table 4.

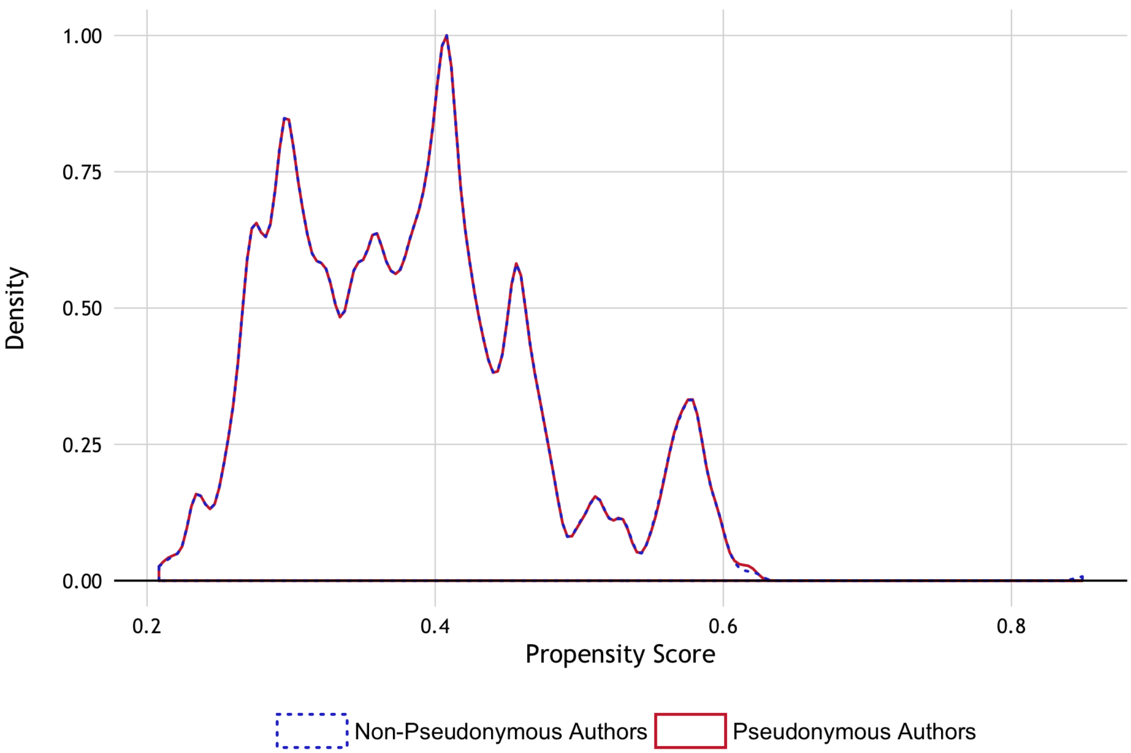


Figure 8: Log Volume (Pseudonymous − Non-Pseudonymous Difference)

This figure plots, on the y-axis, the average difference in log volume between pseudonymous and non-pseudonymous articles, and on the x-axis, the day relative to the date of publication. The blue dashed line plots the average pseudonymous vs. non-pseudonymous difference in open interest for call options and the red solid line plots the average pseudonymous vs. non-pseudonymous difference in volume for put options. Both are the residuals from a fixed effect specification, i.e., after subtracting the average pseudonymous vs. non-pseudonymous difference in log volume for calls and puts written on each firm-article in the interval $[t_0 - 9, t_0 + 5]$.