Nonlocal Disadvantage: An Examination of Social Media Sentiment

Robert Giannini
Citadel

Paul Irvine
Texas Christian University

Tao Shu*
University of Georgia and HKUST Business School

Twitter posts covering 1,082 firms from November 2008 to June 2011 reveal that sentiment in nonlocal Twitter posts is negatively related to future returns, and this negative relation is due to nonlocal posts favoring overpriced stocks, which earn lower subsequent returns. In contrast, local posts do not exhibit this failing. Since nonlocal posts dominate social media, this result highlights the danger of a naive reliance on social media sentiment. The nonlocal disadvantage is larger for firms without public news and firms with higher information asymmetry, suggesting that richer information constrains the exuberance of nonlocal investors. (*JEL* G11, G14, O35)

Received date January 19, 2016; Accepted date: June 25, 2017 By Editor Wayne Ferson

The use of sentiment extracted from social network information has exploded as an investment tool in recent years. Institutional investors pioneered the use of social network sentiment information but recently several individual investor trading platforms added social network sentiment information for the use of their retail investors. Unfortunately, very little information on the nature or reliability of this information is known and investors are left to their own devices when applying this information to

* We thank Wayne Ferson (editor) and an anonymous referee for their valuable comments that significantly improved this paper. We thank Stocktwits.com for providing the stock tweets data; Kenneth Ahern, Jonathan Berk, Fangjian Fu, Mark Grinblatt, Bige Kahraman, Harold Mulherin, Jeffry Netter, Mark Seasholes, Sheridan Titman, and Eric Yeung for helpful comments; and seminar participants at the 2014 European Finance Association Meeting, 2014 Utah Winter Finance Conference, the 2013 FIRS Conference, the 2013 Asian Finance Association Conference, the 2012 China International Conference of Finance, the 3rd TCFA Best Paper Symposium, and University of Georgia for helpful comments. Shu appreciates the financial support from the Terry Sanford Award at the University of Georgia. A substantial amount of the work for this paper was done when Shu was a visiting associate professor at HKUST. A previous version of the paper was titled "Do local investors know more? A direct examination of individual investors' information set." Send correspondence to Tao Shu, Department of Finance, Terry College of Business, University of Georgia, Athens, GA 30602; telephone: 706-542-4450. E-mail: taoshu@terry.uga.edu.

Published by Oxford University Press on behalf of The Society for Financial Studies 2017.
This work is written by US Government employees and is in the public domain in the US.
doi:10.1093/rapstu/rax020 Advance Access publication July 31, 2017

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

their investment decisions.[1] For example, some observers are puzzled by the fact that there is an active market for Twitter sentiment, which is based on an open source stream of 140 character bits of information (Eggers 2014).[2]

This paper evaluates the validity of social network information using a novel approach to investigate whether local and nonlocal social network sentiment can predict stock returns. Whereas previous studies of local advantage such as Ivkovic and Weisbenner (2005) and Seasholes and Zhu (2010) examine the returns to investments by individual investors, we use a large sample of Twitter posts to examine investors' sentiment about local and nonlocal firms. Because of Twitter's popularity and influence, Twitter has a growing impact on the financial markets. For example, in April 2013 the U.S. Securities and Exchange Commission approved using Twitter to communicate company announcements.

We obtain a unique data set containing 202,616 Twitter posts from 1,944 Twitter users, covering 1,082 publicly traded U.S. companies from November 11, 2008 to June 10, 2011. We begin by examining the overall ability of Twitter messages to predict stock returns. Specifically, we classify the positive, negative, or neutral sentiment in a Twitter post using a maximum likelihood classification (ME) approach and then aggregate the sentiment of all posts on a stock. The ME classification technique does more than simple algorithms that just count the numbers of positive or negative words, as it allows both individual words and phrases, as well as their strength, to indicate sentiment.[3] Further tests on abnormal returns suggest that the ME approach, compared with the naive Bayesian method currently standard in the literature, can better identify sentiment.

Since social network sentiment can differ across posters in identifiable ways, we classify Twitter posts into local and nonlocal posts according to the distance between a user's location and corporate headquarters. Although existing studies provide strong evidence of local advantage for institutional investors (e.g., Coval and Moskowitz 2001; Baik, Kang, and Kim 2010; Bernile, Kumar, and Sulaeman 2015), the empirical evidence is mixed for individual investors.[4] Ivkovic and Weisbenner (2005) examine a sample of

1 For example, the broker TDAmeritrade provides aggregate sentiment levels and twitter post volume on individual stocks.

2 As reported in a *Wall Street Journal* article (Dwoskin 2013), clients of Twitter's data business include hedge funds and other investors.

3 It is difficult to compare the accuracy of ME classification with previous studies in the finance literature because they, in general, use key word counts in the empirical analysis without examining the proportions of correct and incorrect identifications of sentiment. Loughran and McDonald (2011) report that 73.8% of the negative word counts based on the Harvard Dictionary are not associated with a negative meaning in a financial context. Their sample is 10-K reports, rather than Internet messages.

4 Researchers also have documented local advantage for various financial market participants including analysts (Malloy 2005; Bae, Stulz, and Tan 2008), commercial and investment banks (Butler 2008; Agarwal and Hauswald 2010), institutional shareholders as monitors (Gaspar and Massa 2007; Ayers, Ramalingegowda, and Yeung 2011), and acquirers (Almazan et al. 2010).

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

34,517 households from 1991 to 1996. They find that individual investors earn an abnormal return of 3.2% per annum on their local holdings relative to their nonlocal holdings. However, Seasholes and Zhu (2010) use a calendar-time portfolio approach to show that individual investors earn no significant alpha on their local holdings. Seasholes and Zhu conclude that "individuals do not seem to have value-relevant information about the local stocks they trade."

We first examine the full sample of posters and find that the sentiment in their posts on average *negatively* predicts stock returns. For example, a one-standard-deviation increase in the 2-week sentiment measure is associated with an 8.3-basis-point (bp) *decrease* (t-stat $= -5.01$) in abnormal stock returns in the subsequent week. This result is robust for return windows from 2 days to 1 month and for alternative sentiment measurement windows. We then classify local and nonlocal posts according to whether the distance between a user's location and corporate headquarters is within 100 miles, and examine whether the opinions of local posters, compared with nonlocal posters, better predict stock returns.[5] We find that nonlocal sentiment exhibits strong negative return predictive ability. For example, a one-standard-deviation increase in nonlocal sentiment predicts a decrease of 16 bps (t-stat $= -5.20$) in weekly stock returns. In contrast, local sentiment has no significant predictive ability. These results reveal that for individual posters, there is a "nonlocal disadvantage" instead of "local advantage." The nonlocal disadvantage, measured as the differential return predictive ability between local and nonlocal sentiment, is a statistically significant 15 bps (t-stat $= 3.69$) per week.

The nonlocal disadvantage persists for return measurement windows from 2 days to 1 month, and remains robust to controls for priced factors and firm characteristics, alternative definitions of local posts, alternative samples of posters, and alternative approaches to measuring sentiment in the posts.[6]

We further examine long-short strategies based on Twitter sentiment. We first examine a strategy that goes short stocks with favorable Twitter sentiment and goes long stocks with unfavorable Twitter sentiment, and find that the contrarian predictive ability of Twitter sentiment is accounted for by the abnormally low return of the short portfolio, namely those stocks with favorable Twitter sentiment.

We then examine a strategy that goes short the portfolio containing firms with more favorable nonlocal sentiment than local sentiment (nonlocal

5 We conduct robustness tests by classifying local and nonlocal posts using the 50-mile criterion (Ivkovic and Weisbenner 2007) instead of the 100-mile criterion or whether the Twitter user and the company are located in the same state. The results hold with both alternative approaches.

6 To address the concern of potential noise in the Twitter posts, we use a subsample of users on Stocktwits.com's list of "recommended" contributors who generally have a large following, a long track record, post meaningful comments, and are notable within the social network. Nonlocal disadvantage persists when we restrict the sample to recommended users.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

favorable stocks) and goes long the portfolio containing firms with less favorable nonlocal sentiment than local sentiment (nonlocal unfavorable stocks). Consistent with nonlocal disadvantage, the abnormal profit is around 6 bps per day and is statistically significant. This profit also comes from the abnormally low return of the nonlocal favorable portfolio. Additionally, the strategy is more profitable when we exclude firms with less information asymmetry or limits to arbitrage, due to the larger negative return of the short portfolio. These results indicate that nonlocal disadvantage is caused by higher nonlocal sentiment in overpriced stocks, which earn lower subsequent returns.

Why is nonlocal sentiment higher in overpriced stocks? It is possible that nonlocal individuals who lack local knowledge are attracted to "glamor" stocks, which tend to be overpriced. To explore this possibility, we examine the relation of Twitter sentiment with contemporaneous return and trading volume. We find that compared to local sentiment, nonlocal sentiment has much stronger positive relations with contemporaneous abnormal trading volume and contemporaneous return, especially among high volatility stocks where overpricing is more prevalent.

The finding of nonlocal disadvantage suggests that nonlocal investors, who suffer more from information asymmetry between firms and investors, make consistent mistakes in predicting stock returns. Therefore, we examine how the nonlocal disadvantage varies across alternative information environments. We find that nonlocal disadvantage is much larger in the firms without news coverage (45 bps per week) than in the firms with news coverage (10 bps per week). This finding is consistent with Tetlock's (2010) contention that public news releases can level the playing field for uninformed investors. Additional results also show a significantly positive association between nonlocal disadvantage and various proxies for information asymmetry.

We conduct two further analyses that examine the relation between nonlocal disadvantage and stock mispricing. The first analysis is based on Stambaugh, Yu, and Yuan (2015), who show that higher idiosyncratic volatility indicates greater limits to arbitrage and therefore greater mispricing, particularly overpricing. We find that nonlocal disadvantage is large and significant in high volatility stocks but small and insignificant in low volatility stocks, consistent with nonlocal disadvantage being associated with overpricing. Our second test is based on La Porta et al. (1997), who suggest that if an anomaly is associated with mispricing, then it will be stronger in the earnings announcement window, as the release of earnings helps correct mispricing.[7] We repeat the return regressions using the 2-day earnings announcement return as the independent variable. We find that, consistent with this mispricing prediction, the nonlocal disadvantage with respect to the

[7] Engelberg, McLean, and Pontiff (2016) use this approach to study a strategy that combines 94 anomalies and find that the anomalies are much stronger on earnings announcement days.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

2-day earnings announcement return is 0.478% (t-stat $= 2.78$), about 7.5 times as large as the nonlocal disadvantage of 0.064% from our baseline regressions of 2-day returns.

Finally, we examine whether nonlocal individuals are also more affected by the biases in analyst forecasts and recommendations. We do this to extend the existing literature that investor characteristics or experience can lead to amelioration of investor's behavioral biases.[8] Our test is motivated by Malmendier and Shanthikumar (2007, 2014), who find that small traders do not recognize that analysts issue overly optimistic recommendations and forecasts and therefore react to analysts' opinions naively. We first examine how nonlocal posters are affected by analyst optimism, defined as the difference between consensus analyst forecast and actual earnings. We find that nonlocal sentiment has a significantly positive relation with contemporaneous analyst optimism, whereas local sentiment has no significant relation with analyst optimism. Next, we examine how nonlocal posters are affected by consensus analyst recommendations. We find that both nonlocal sentiment and local sentiment are higher in buy recommendation stocks, but this relation is significantly stronger for nonlocal sentiment than for local sentiment. In contrast, local sentiment and nonlocal sentiment respond similarly to the more informative sell recommendations.

We show that nonlocal investors' sentiment is more susceptible to naive reliance on the opinions of financial analysts. As naive reliance on analysts' consensus opinions underperforms in our sample period, the ability of nonlocal investors to increase their reliance on local investor opinion has the potential to reduce their errors.[9] This evidence suggests that local investors appear to use their own local knowledge as a substitute for analysts' opinions, which allows them to overcome the pitfalls sometimes associated with naive reliance on analysts' opinions. These tests highlight the potential for crowdsourcing sites, such as Stocktwits.com, to improve individual investor performance and extend our knowledge of how individual investors can attenuate the natural biases of human decision-making.

Twitter posts reflect posters' sentiment, but not necessarily their trading decisions. This feature avoids the impact of various trading or capital constraints, but it does not allow us to observe the locals and nonlocals' investing decisions or trading performance. Our results based on social media sentiment complement the existing literature on individual investor's local

[8] For example, Feng and Seasholes (2005) examine Chinese investors and contend that the disposition effect can be largely eliminated through experience and sophistication. Seru, Shumway, and Stoffman (2010) counter that improvements in trading performance attributed to experience are instead caused by self-selection when some traders learn they are of low trading ability and discontinue active trading. Korniotis and Kumar (2011) suggest that better investment decisions are correlated with increasing age only to the point at which the decreasing cognitive ability of older people reduces the sophistication of their investment decisions.

[9] Jegadeesh et al. (2004) warn that naive reliance on the level of *consensus* recommendation can sometimes mislead investors, as it does in our sample. The stronger nonlocal reaction suggests a greater reliance on these sometimes misleading recommendations.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

advantage based on their actual trading. On one hand, consistent Ivkovic and Weisbenner's (2005) portfolio tests, we find that nonlocal sentiment has a more negative association with future return than local sentiment. On the other hand, consistent with Seasholes and Zhu's (2010) conclusion, we find that even local sentiment is not positively associated with future return.

Our study contributes to the literature on social media and financial markets. Motivated by the rapid growth in social media in the past two decades, financial researchers have started to investigate whether stock-specific internet messages contain value-relevant information or just noise. Several studies find that messages posted on internet stock message boards (e.g., Yahoo! Finance) have little to no predictive power for stock returns, suggesting that these messages may simply contain noise (Tumarkin and Whitelaw 2001; Antweiler and Frank 2004; Das, Asis, and Tufano 2005). We extend these studies to mainstream social media and, in contrast to these studies, document that sentiment in Twitter posts negatively predicts stock returns. We also report that messages are not homogeneous and certain message characteristics, such as the geographical location of posters can eliminate these errors. Our results provide evidence that social media messages can be effectively used by individual investors, but that profitable investing strategies must be more sophisticated than blindly following the current sentiment information.[10]

1. Data and Sample Selection

1.1 Twitter and financial markets

Twitter is a micro-blogging application whereby users are able to post short thoughts of no more than 140 characters, called tweets. Twitter was started as a social network, but its worldwide popularity and broad user base have led to its fast growing impact on many aspects of people's lives, including, for example, presidential politics. Twitter also has been related to the financial markets. Paul Hawtin, founder of Twitter hedge fund Derwent Capital, claims "Today, social media creates a vast amount of information and it has been proven that the sentiment derived from it can predict stock market movements." In April 2013, the U.S. Securities and Exchange Commission approved using Twitter to communicate company announcements. On April 24, 2013, the Dow Jones industrial average immediately plunged by more than 140 points after a hacker sent out a false tweet from Associated Press's account.

We collect Twitter posts from Stocktwits.com, an open micro-blogging site powered by Twitter with a focus on financial markets. Stocktwits.com was founded in 2008 and has since become a popular Web site for Twitter users to exchange investment information. Since its inception, Stocktwits.com has

[10] Chen et al. (2014) provide evidence supporting the value relevance of particular Internet sites.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

been covered by major news media such as The New York Times and CNNMoney.com. In 2010, Stocktwits was named Time.com's top 50 best Web sites as well as Fast Company's top 10 innovative companies in finance.

1.2 Collection of Twitter posts and sample construction

Figure 1 provides an example of the stream of Twitter posts that comprise our sample. Twitter users comment about a specific company by referring to the company's ticker preceded by a "$" hashtag, for example, "$MSFT and $AAPL are a buy!" Hashtagging allows us to extract specific company references.[11]

Our sample contains the twitter posts about publicly traded companies from Stocktwits.com from November 11, 2008 to June 10, 2011. The sample includes two parts. The first part consists of the posts from users in Stocktwits.com's list of "recommended" contributors from November 11, 2008 to July 10, 2009. The "recommended" users generally have a large following, a long track record, post meaningful or interesting comments, and are influential within the social network. We directly collected these tweets from Stocktwits.com. The second part consists of posts from all users of Stocktwits.com from July 11, 2009 to June 10, 2011, which we obtained from the company. For each post in the data, we have the content of the post, the associated ticker symbol(s), the date and the time of the post, and the blogger's account ID and number of followers.

We require the following information for a tweet to be included in our sample. First, since some of the symbols represent non-stock assets such as gold, foreign currencies, or indices, we identify stock tweets by matching to stock tickers in CRSP, and then match stock tickers to PERMNOs, which is the unique firm identifier in our analysis. Next, we collect a Twitter user's location on the user's profile page on Stocktwits.com by searching the account ID and require users to live in the continental U.S. and provide location information at the city (county) level because we require both the state and the city information to calculate the distance to corporate headquarters. Appendix A.1 and A.2 provide details about these steps.

We further require the sample firms to have available CRSP data, have headquarters located in the contiguous United States, and have at least ten Twitter posts and one local Twitter post over our sample period.[12] We drop penny stocks that are priced below two dollars at the end of the

[11] In the case of multiple company references in one post, like in the example above, each reference is counted as a unique post. Among the original posts, 88% cover only one symbol, 7% cover two symbols, and 5% cover more than two symbols.

[12] Our results are robust when we require the firm to have at least one Twitter post over the sample period or when we do not require the firm to have at least one local Twitter post over the sample period.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

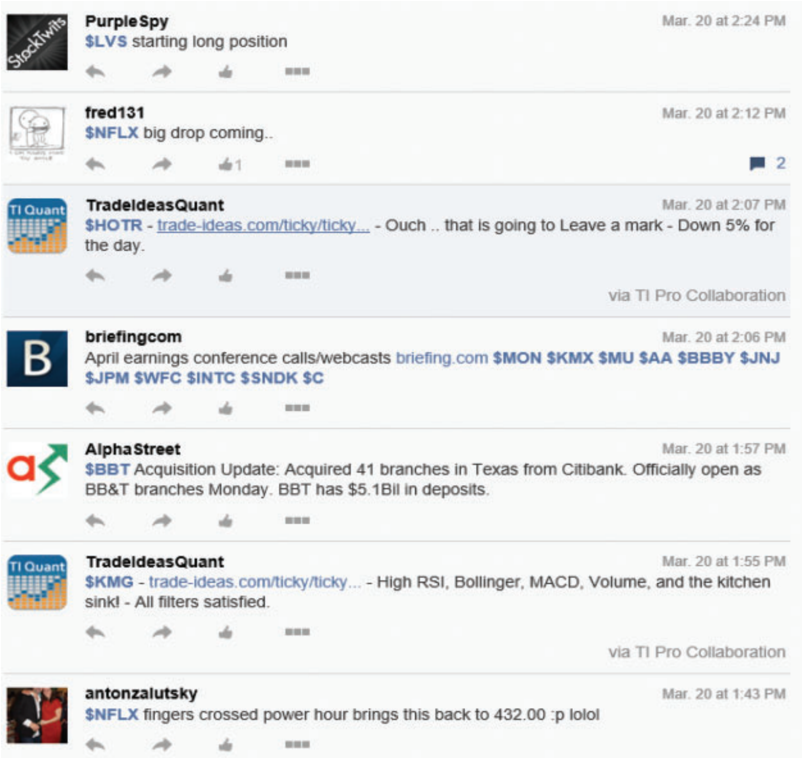

Figure 1
Example of Twitter stream
This figure shows the interface that a Stocktwits.com user would see. Company tickers can be seen after the $ hashtags.

previous year.[13] Our final sample contains 202,616 Twitter posts from 1,944 Twitter users, covering 1,082 publicly traded U.S. companies from November 11, 2008 to June 10, 2011. This sample includes two subsamples as mentioned above. The first subsample consists of 22,368 posts from 45 "recommended" users from November 11, 2008 to July 10, 2009, and the second subsample consists of 180,248 posts from all users of Stocktwits.com from July 11, 2009 to June 10, 2011. We repeat the analysis in this paper for the two subsamples and the results are similar in terms of both economic and statistical significance.

Figure 2 plots the locations of the sample Twitter users. Users live in all of the states in the contiguous United States, except North Dakota. The highest percentages of users are in the states of New York (18%), California (17%), Illinois (9%), Texas (7%), and Florida (7%). The remaining users reside in 42

13 Our results are similar when we include penny stocks in the sample.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

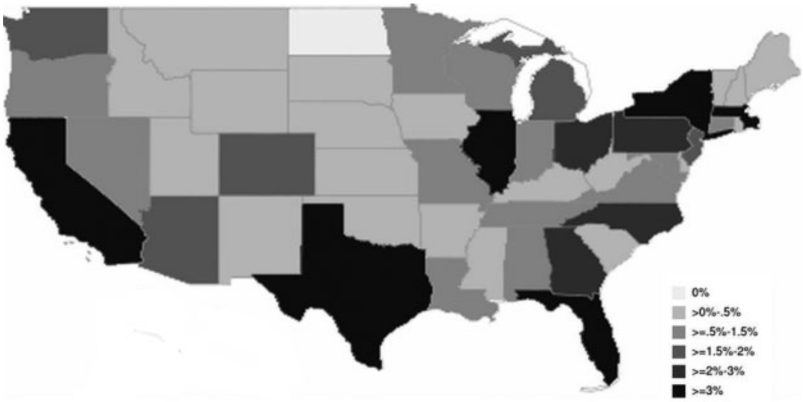

Figure 2
Geographical distribution of sample Twitter users
This figure plots the geographical distribution of the Twitter users in our sample. We divide the states with six groups depending on the percentages of sample users: 0%; between 0% and 0.5%; between 0.5% and 1.5%; between 1.5% and 2%; between 2% and 3%; greater than 3%. The states with higher percentages are marked with darker shades.

other states and Washington DC, with no other state accounting for over 5% of sample users. The geographical distribution of Twitter users is consistent with the distributions of U.S. population and economic activity. This dispersed distribution avoids the issue of geographical clustering.[14]

We obtain accounting data for our sample firms from Compustat, and data on analyst coverage, analyst forecasts, and analyst recommendations from the I/B/E/S summary file. For the construction of abnormal returns, we obtain the daily returns of the three Fama-French factors and the momentum factor (UMD) from Kenneth French's data library.[15] Some of our tests use news articles collected from Factiva, and we will describe the details of the construction of news data when we discuss the corresponding tests.

1.3 Classifying local and nonlocal posts

We calculate the straight line geographic distance between the location of each Twitter user and the headquarters of each company in our sample using longitude and latitude coordinates. We assign longitude and latitude coordinates to the user locations according to the state and city information on their profiles. We further obtain ZIP codes of corporate headquarters from Compustat and assign the corresponding longitude and latitude

[14] Geographic clustering is a common problem in U.S. financial research as many financial intermediaries are clustered in the New York City area (see Anand et al. 2011).

[15] We thank Professor Kenneth French for making these data available.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

coordinates.[16] We then calculate the distance between Twitter user and company headquarters using the following equation:

$$Distance = 7921 * arcsin(sqrt((sin\,((0.017 * lat2 - 0.017 * lat1)/2))^2 + cos\,(0.017 * lat1) * cos\,(0.017 * lat2) * (sin\,((0.017 * long2 – 0.017 * long1)/2))^2)), \quad (1)$$

where *lat1* and *long1* are the latitude and longitude coordinates of a Twitter user and *lat2* and *long2* are the latitude and longitude coordinates of corporate headquarters.[17]

We classify a Twitter post as local (nonlocal) if the distance between the Twitter user and the corporate headquarters is within (more than) 100 miles. Previous studies use various criteria of distance to classify local stocks, from 62 to 250 miles (e.g., Coval and Moskowitz 2001; Ivkovic and Weisbenner 2005; Malloy 2005; Seasholes and Zhu 2010). We adopt the moderate 100-mile criterion and classify 23,918 posts as local and 178,698 posts as nonlocal. For robustness, we also try alternative criteria and discuss the results in Section 2.4.

1.4 Qualifying the sentiment in twitter posts

The unique features of Stocktwits posts, such as addressing replies, and the 140 character restriction produce a language that is markedly different from standard English. Conventional word counts using standard English dictionaries are unlikely to be useful in interpreting Stocktwits posts.

We use the maximum entropy (ME) approach, which endogenously creates a dictionary of terms, to classify the sentiment in Twitter posts. The ME approach derives sentiment from the statements in posts by applying a maximum likelihood algorithm to the data. The information in Twitter posts can be subtle. For example, the statement "You would be crazy to sell $GOOG right now" contains the word "sell," which we would unconditionally assume has a negative connotation. However, the statement "crazy to sell" is a positive statement. ME classification is considered the most robust technique for information classification because it controls for the conditional dependence of words (Pang, Lee, and Vaithyanathan 2002). Unlike the less sophisticated procedures that handle each word as an unconditional feature, ME classification uses the information contained in multiple word phrases, such as "crazy to sell," to more accurately classify sentiment.

[16] We match longitude and latitude coordinates to ZIP codes using the database from http://www.getzipcodedata.com/#.

[17] This equation is provided by SAS at http://www2.sas.com/proceedings/sugi31/143-31.pdf. This approach is based on the great circle distance model that is similar to the distance equations used in the literature (e.g., Ivkovic and Weisbenner 2007), but Equation (1) provides greater accuracy at small distances. More details about the distance models can be found at http://en.wikipedia.org/wiki/Great-circle_distance.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

In addition to controlling for the conditional dependence of words, the ME classification also avoids the misidentification issue associated with alternative approaches that simply rely on key-word frequencies. For example, Loughran and McDonald (2011) show that in the textual analysis of 10-K reports, almost three-fourths (73.8%) of the negative word counts according to the widely used Harvard Dictionary are attributable to words that are typically not negative in a financial context (e.g., tax, cost, capital, board, and liability). Other words on the Harvard list (e.g., mine, cancer, crude, tire, and capital) are more likely to identify a specific industry segment than reveal a negative financial event. ME classification does not suffer the noise introduced by key-word selection because the identification is based on a large training sample of Twitter posts that we hand classify.[18,19]

The general idea of ME classification is that when nothing is known about a distribution, the distribution should be uniform, that is, have maximum entropy. Consider the example of trying to classify a document as positive, negative, or neutral, where we are only told that 50% of documents that contain the word "buy" are considered positive. Intuition tells us that if the document has the word "buy" in it, then there is a 50% chance that it is a positive post, a 25% chance of being negative, and a 25% chance of being neutral. If our document did not have the word "buy" in it, then we would just assume an equal distribution of a 33% chance that the document falls into each category. Thus, if we knew nothing about our document, we begin with a uniform distribution with equal likelihood for each sentiment category. This is the essence of ME classification. In practice, this process is constrained by many features, and the calculations for conditional probabilities become complex, but the logic is still the same as our simple example.[20] We formally describe the ME procedure in Appendix A.3.

The Twitter data is generally marketed and used in the industry to gauge the strength and momentum of intraday price moves. Specifically the data is employed to show whether strong investor sentiment confirms intraday price moves. To test whether investor sentiment actually confirms intraday price moves, we estimate a regression of daily abnormal return on the same-day Twitter sentiment measured as sum of the sentiment of Twitter posts on the day (results discussed in Section 4.2). We find a significant positive relation between Twitter sentiment and same-day return. This result suggests that

[18] Additionally, many previous studies using the Harvard list only count negative words because they find little incremental information in Harvard's positive word list (e.g., Tetlock 2007; Engelberg 2008). In contrast, the ME classification is based on both positive and negative words found in the messages.

[19] ur approach is similar in spirit to that of Loughran and McDonald (2011), who exogenously define a dictionary that is suitable to particular types of financial information and then use that dictionary to evaluate a large data set. We exogenously classify sentiment in a training data set and allow the ME program to determine the likelihood that a particular word or phrase represents a particular sentiment.

[20] Thomson analytics uses a related method to assign sentiment to Reuters' news articles (Hendershott, Livdan, and Schürhoff 2015).

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Twitter data can be useful in improving the profitability of intraday trading strategies.

We also try classifying the sentiment of Twitter posts using the naive Bayesian (NB) approach proposed by the existing literature (e.g., Li 2010; Huang, Zang, and Zheng 2014). We conduct robustness tests using the NB approach and find similar results on nonlocal disadvantage. Section 2.5 discusses the NB approach and the corresponding results.

2. Does Local and Nonlocal Sentiment Predict Stock Returns?

2.1 Summary statistics

Panel A of Table 1 summarizes the characteristics of our sample. A typical firm in our sample has a market capitalization of $8,046 million, a book-to-market ratio of 0.59, and is followed by 10.73 analysts. For a comparison, an average firm in the contemporaneous CRSP universe has a market cap of $2,431 million, a book-to-market ratio of 0.79, and is followed by 4.03 analysts. Since we require sample firms to have at least ten Twitter posts during the sample period, these comparisons suggest that the firms covered by Twitter users tend to have larger size, higher analyst coverage, and lower book-to-market ratio.[21] We also report average idiosyncratic volatility and daily return for sample firms. Idiosyncratic volatility for a firm-day is the standard error of residuals from the time-series regressions of a firm's excess returns on the daily market factor (MKT) in the one-year window up to the end of previous month.[22] A typical sample firm has idiosyncratic volatility of 0.032 and average daily return of 17 bps, similar to the contemporaneous CRSP universe.

Panel B of Table 1 presents summary statistics of the sample Twitter posts. The average sentiment of Twitter posts is positive, which is consistent with the recovery of the stock market during the sample period. Although both the local and nonlocal tweets are positive, local tweets are less so than nonlocal tweets. Regarding the Twitter coverage, a sample firm receives on average 187 posts during our sample period, including 165 nonlocal posts and 22 local posts.

To examine the why certain stocks receive interest on Twitter, we present in panel C of Table 1 the firm-level regression of Twitter coverage on firm characteristics and stock market metrics. In the first model, the dependent variable is the total number of Twitter posts for a firm during the sample period. The independent variables include average firm and stock metrics during the sample period, including market capitalization, book-to-market ratio, analyst coverage, idiosyncratic volatility, and average daily return.

[21] We drop penny stocks priced below $2, and doing so makes our sample firms larger than those in the CRSP universe.

[22] We require at least 100 daily return observations in the estimation window.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 1
Summary statistics

A. Characteristics of sample firms

	Mean	SD	P10	P25	P50	P75	P90
Market capitalization ($M)	8,046	23,570	209	480	1,528	5,099	17,730
Book/market ratio	0.59	0.55	0.12	0.26	0.48	0.80	1.21
Analyst coverage	10.73	7.39	1.98	4.90	9.65	15.66	20.77
Idiosyncratic volatility	0.032	0.017	0.016	0.022	0.029	0.039	0.051
Daily stock return (%)	0.17	0.27	0.03	0.09	0.016	0.25	0.35

B. Summary statistics of sample Twitter posts

	Mean	SD	P10	P25	P50	P75	P90
Sentiment of posts	0.31	0.68	−1.0	0.0	0.0	1.0	1.0
Sentiment of local posts	0.22	0.67	−1.0	0.0	0.0	1.0	1.0
Sentiment of nonlocal posts	0.32	0.68	−1.0	0.0	0.0	1.0	1.0
# posts per firm	187.26	735.35	16.0	27.0	61.0	134.0	302.0
# local posts per firm	22.11	129.61	1.0	2.0	5.0	13.0	34.0
# nonlocal posts per firm	165.15	620.39	11.0	22.0	52.0	118.0	284.0

C. Regressions of Twitter coverage on firm characteristics

	Dependent variable		
	#Twitter posts	#local posts	#nonlocal posts
ln(ME)	−0.108*	−0.124**	−0.085
	(−1.84)	(−2.07)	(−0.25)
Book-to-market ratio	−0.052*	−0.020	−0.102*
	(−1.78)	(−0.66)	(−1.74)
Ln(Analyst Coverage)	0.069*	0.007	0.080*
	(1.68)	(0.16)	(1.96)
# news articles	0.445***	0.443***	0.435***
	(9.91)	(9.67)	(9.70)
Idiosyncratic volatility	0.044	0.007	0.050
	(1.18)	(0.23)	(1.35)
Ave. daily ret.	0.042	0.018	0.046
	(1.42)	(0.58)	(1.56)
Adjusted R^2	0.154	0.124	0.154
Number of obs	1,082	1,082	1,082

Panel A reports summary statistics for the 1,082 firms in our sample from November 11, 2008 to June 10, 2011. For a firm-day, market capitalization is measured at the end of previous year. Book-to-market ratio is book equity divided by market capitalization measured at the end of fiscal year. A firm's book-to-market ratio of fiscal year ending in calendar year t is matched to firm-days from July of t+1 to June of t+2. Book-to-market ratios are winsorized at 1% and 99% cutoff points. For a firm-day, analyst coverage is the number of analysts covering the firm in the previous month; idiosyncratic volatility the standard deviation of residuals from time-series regression of the firm's excess returns on the market excess returns (MKT) in the one-year window ending in previous month; and daily return is the daily raw return. Idiosyncratic volatilities are winsorized at the 99% cutoff points. We first calculate the average of firm characteristics for a firm across the firm-days and then report the distribution of average characteristics across firms. Panel B presents the summary statistics of the sample Twitter posts, including the sentiment of a post identified using the maximum entropy (ME) approach, and the number of Twitter posts for a sample firm during the sample period. A Twitter user is local (nonlocal) to a firm if the user's location is less than (more than) 100 miles from the firm's headquarters. Panel C presents firm-level cross-sectional regressions of Twitter coverage on firm characteristics. The independent variable is the number of Twitter posts, local Twitter posts, or nonlocal Twitter posts for a firm during the sample period. The dependent variables include firm characteristics measured like in panel A. *ln(ME)* is natural log of market capitalization. *ln(Coverage)* is natural log of 1 plus analyst coverage. *# news articles* is the total number of news articles covering a firm during the sample period. We standardize both the dependent variables and independent variables. We also repeat the regression using alternative dependent variables including the number of local posts and the number of nonlocal posts. The regressions include a constant term that is not reported for brevity. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Because of the existing literature on "attention attracting" for individual investors (e.g., Barber and Odean 2008; Hirshleifer, Lim, and Teoh 2011), we also include as an independent variable the total number of news articles for a firm during the sample period (the collection of news article is discussed in Section 3.1). The coefficients on the book-to-market ratio and firm size are significantly negative but those on analyst coverage and news articles are significantly positive. These results suggest that Twitter coverage is higher for growth firms, smaller firms, and firms with more news coverage and analyst coverage. Since we standardize the dependent and independent variables, the results suggest that news coverage has the largest influence on Twitter coverage among all the independent variables. We also repeat the regressions for local and nonlocal Twitter coverage respectively. Both local coverage and nonlocal coverage are strongly affected by news. Regarding firm characteristics, local coverage is less sensitive to book-to-market ratio and analyst coverage than nonlocal coverage, but is more tilted towards smaller firms than nonlocal coverage.

2.2 Stock return predictive ability using all tweets

Before the examination of locals versus nonlocals, we first examine the overall stock return predictive ability of Twitter sentiment. Specifically, we estimate the following daily panel regression:

$$CAR[t,\ t+k]_i = \alpha_0 Sent_{it} + \sum \beta_j AR_{it-j} + \sum \gamma_i D_i + \varepsilon_{it}, \qquad (2)$$

where $CAR\ [t,\ t{+}k]_i$ is cumulative abnormal returns of firm i from day t to $t{+}k$. For our tests, we examine abnormal returns in 2- ($k = 1$), 5- (k=4), 10- ($k = 9$) and 20-day (k=19) windows. We follow the spirit of the classic event-study (e.g., Fama et al. 1969) but calculate daily abnormal return as residuals from a four-factor model. To calculate the abnormal return for firm i on day t, we estimate a four-factor model of firm i's daily returns in the [t-150, t-31] window. The model includes the three daily Fama-French (MKT, SMB, HML) factors and the daily momentum factor (UMD). We then use the factor loadings from this regression to calculate the abnormal return, AR_{it}, on day t.[23]

The independent variable is the sum of local and nonlocal sentiment. Specifically, local sentiment ($Local\_Sent_{it}$) is the aggregate sentiment of local Twitter users for firm i over the 2-week period prior to day t. We first assign the scores of -1 (negative), 0 (neutral), or 1 (positive) to each local Twitter post about firm i in the 2 weeks prior to day t using the ME classification techniques described in Section 1 and then sum the scores. We assign zero to the sentiment measure if a firm is not covered by any local Twitter post in the 2-week period. For robustness, we also repeat the analysis including only the

[23] We require at least 30 daily return observations in the estimation window.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

firm-days with at least one Twitter post in the sentiment measurement period, and discuss the results in Section 2.4. We construct nonlocal sentiment ($NonLocal\_Sent_{it}$) as the aggregate sentiment of nonlocal Twitter users for firm i over the 2-week period prior to day t. We then standardize both the local and nonlocal sentiment to have a mean of zero and standard deviation of one, and calculate total sentiment ($Sent_{it}$) as the sum of local and nonlocal sentiment. We further include firm fixed effects (D_i) to control for firm-specific characteristics, and ten lags of daily returns ($AR_{it\text{-}j}$) to control for short-term return reversals and microstructure effects. We calculate t-statistics using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations.[24]

Panel A of Table 2 presents the results of the return regressions in equation (2). Notably, we observe that the coefficients on investor sentiment are significantly negative for all of the return windows. For example, in the model of 5-day return, the estimated coefficient on investor sentiment is −0.083 (t-stat = −5.01), indicating that a one-standard-deviation increase in the sentiment measure is associated with an 8.3-bp *decrease* in subsequent weekly return. Panel B repeats the regressions, but with sentiment in the 1-week period prior to return measurement instead of the 2-week period, and the negative return predictive ability persists for all return windows. This finding illustrates that one may lose money in the stock market by simply following the opinions of Twitter posts.

2.3 Nonlocal disadvantage

In this section, we compare the predictive ability of local Twitter users and their nonlocal peers by estimating the following daily panel regression:

$$CAR[t,\ \ t+k]_i = \alpha_1 Local\_Sent_{it} + \alpha_2 NonLocal\_Sent_{it} + \sum \beta_j AR_{it-j} + \sum \gamma_i D_i + \varepsilon_{it}, \tag{3}$$

which is similar to Equation (2), except that we examine local and nonlocal sentiment separately. To ease the assessment of economic significance, we standardize the local and nonlocal sentiment. The coefficients α_1 and α_2 indicate the predictive ability of local and nonlocal posters.

We report the results in the panel A of Table 3. Nonlocal investors have the same significantly negative relation between sentiment and return previously reported in Table 2 for the combined sample. For example, in the model of 5-day returns, the estimated coefficient is −0.163 (t-stat −5.20) on nonlocal sentiment. In contrast, the negative relation between sentiment and future return disappears for local posters, with the coefficient on local sentiment being only −0.015 (t-stat = −0.76). The difference between locals and

[24] The Driscoll-Kraay standard errors are similar in spirit to the Newey-West standard errors, but the former corrects both time-series and cross-sectional correlations in the panel regression setting.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 2
Panel regressions of abnormal stock returns on investor sentiment

	Dependent variable			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
A. Regressions on 2-week sentiment				
2-week sentiment	−0.036***	−0.083***	−0.159***	−0.312***
	(−5.04)	(−5.01)	(−5.20)	(−6.14)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
B. Regressions on 1-week sentiment				
1-week sentiment	−0.035***	−0.080***	−0.145***	−0.281***
	(−5.45)	(−5.44)	(−5.65)	(−6.41)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082

Panel A presents panel regressions of abnormal stock returns on prior investor sentiment. The dependent variable is 2-, 5-, 10-, or 20-day cumulative abnormal returns (measured in percentage). To calculate daily abnormal return for a firm-day, we first estimate a Fama-French four-factor regression for the firm in the previous 150-day rolling window and then use the estimated factor loadings to calculate abnormal returns for the firm-day. The independent variables include investor sentiment in the 2-week windows prior to return measurement. Investor sentiment is the sum of the standardized local sentiment and nonlocal sentiment in the 2 weeks prior to return measurement. All regressions include firm fixed effects with lagged daily returns in the previous 10 trading days as controls. Panel B is similar to panel A, but includes sentiment in the lagged 1-week window instead of lagged 2-week window. *t*-statistics (reported in parentheses) are calculated using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

nonlocals, or the nonlocal disadvantage, is 14.8 bps in weekly returns (t-stat $= 3.69$). This nonlocal disadvantage is both economically and statistically significant. Additionally, nonlocal disadvantage is large and significant for the other return windows of 2, 10, and 20 days. These results provide strong evidence of nonlocal disadvantage among the individual investors in our sample.

2.4 Robustness tests

We conduct various robustness tests on nonlocal disadvantage. First, we repeat the regressions with 1-week lagged sentiment measurement window instead of a 2-week lagged sentiment measurement window in panel B of Table 3 and observe similar results. Second, we construct abnormal returns using benchmark portfolios instead of factor models. Specifically, panel C of Table 3 presents the baseline regressions using DGTW characteristic-adjusted returns (Daniel et al. 1997; Wermers 2004) as dependent variable. The results are similar to the regressions using factor-adjusted returns in terms of both economic and statistical significance.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 3

Panel regressions of stock returns on local and nonlocal sentiment

A. Regressions of Abnormal Returns

	Dependent Variable			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
Local sentiment	−0.006	−0.015	−0.038	−0.076
	(−0.63)	(−0.76)	(−1.14)	(−1.30)
Nonlocal sentiment	−0.070***	−0.163***	−0.299***	−0.585***
	(−5.04)	(−5.20)	(−5.10)	(−5.79)
Local – Nonlocal	0.064***	0.148***	0.261***	0.509***
	(3.36)	(3.69)	(3.76)	(3.92)
Controls of lagged Returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
B. Regressions of abnormal returns: 1-week sentiment				
Local sentiment	−0.014*	−0.020	−0.033	−0.079*
	(−1.65)	(−1.15)	(−1.26)	(−1.73)
Nonlocal sentiment	−0.059***	−0.150***	−0.274***	−0.513***
	(−4.77)	(−5.49)	(−5.66)	(−6.02)
Local – Nonlocal	0.044***	0.130***	0.241***	0.434***
	(2.66)	(3.75)	(4.14)	(4.19)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
C. Regressions of DGTW returns				
Local sentiment	−0.008	−0.015	−0.037	−0.049
	(−0.78)	(−0.72)	(−1.00)	(−0.78)
Nonlocal sentiment	−0.068***	−0.157***	−0.282***	−0.557***
	(−6.14)	(−6.33)	(−5.95)	(−6.98)
Local – Nonlocal	0.060***	0.141***	0.245***	0.508***
	(3.93)	(4.54)	(4.41)	(5.31)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	538,626	538,407	538,626	538,626
Number of PERMNOs	952	952	952	952

Panel A presents panel regressions of stock returns on prior local and nonlocal sentiment. The dependent variable is cumulative 2-, 5-, 10-, or 20-day abnormal returns (measured in percentage). To calculate daily abnormal return for a firm-day, we first estimate a Fama-French four-factor regression for the firm in the previous 150-day rolling window and then use the estimated factor loadings to calculate abnormal returns for the firm-day. The independent variables include local sentiment and nonlocal sentiment in the 2-week window prior to return measurement. To calculate local and nonlocal sentiment, we first classify Twitter posts into local and nonlocal posts according to whether the Twitter users' locations are within 100 miles of the headquarters of the firms mentioned in the posts. We use maximum entropy classification to measure the sentiment of each post and then sum the sentiment measures of the local and nonlocal posts, respectively, in the 2 weeks prior to return measurement. We standardize the independent variables for each regression. For each regression, we further report the difference between the coefficients on local sentiment and nonlocal sentiment. All regressions include firm fixed effects with lagged returns in the previous 10 trading days as controls. *t*-statistics (reported in parentheses) are calculated using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. Panel B is similar to panel A, except that the local and nonlocal sentiment is measured in the lagged 1-week window prior to return measurement. Panel C is similar to panels A, except that the dependent variable is cumulative DGTW-adjusted return, where daily DGTW returns are calculated as a firm's daily raw return minus benchmark return of the benchmark portfolio based on size, book-to-market ratio, and momentum. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Further, we examine whether the results are sensitive to the alternative classification criteria of local posts. We classify local and nonlocal posts according to whether the Twitter users and corporate headquarters are located within 50 miles (panel A of Table 4) or in the same state (panel B of Table 4), and our finding of nonlocal disadvantage persists. For example, when we use the 50-mile criterion in panel A of Table 4, in the 5-day return window the estimated nonlocal disadvantage (α_1- α_2) is 0.139 (t-stat $=$ 3.55), very close to the 0.148 in panel A of Table 3.

When no Twitter post covers a firm during the 2-week sentiment period, we do not drop the observation but treat it as a neutral sentiment by assigning zero to the sentiment measure. For a robustness test, we repeat the regression analysis but include only the firm-days with at least one Twitter post in the sentiment period. Panel C of Table 4 shows that nonlocal disadvantage persists for all return windows.

Drilling down further, we focus on a subgroup of sophisticated users in Stocktwits.com's list of recommended contributors, who generally have a large following, a long track record, post meaningful or interesting comments, and are influential within the social network. There are 101 recommended users in our sample. Panel D of Table 4 shows that nonlocal disadvantage persists when we include only recommended posters. We also repeat the tests by including only posters that have at least five local posts and five nonlocal posts, and nonlocal disadvantage holds (panel E of Table 4).[25]

We also repeat the tests using longer-term returns in the 20-day windows from day 21. Panel F of Table 4 reports the regressions of 20-day abnormal returns in the [21,40], [41,60], [61,80], [81,100], and [101,120] windows. The results show that in the longer-term windows, the coefficients on local sentiment remain insignificant, and the negative coefficient on nonlocal sentiment gradually becomes smaller, leading to a smaller nonlocal disadvantage over time. The nonlocal disadvantage produces a t-statistic of only -1.71 in the [81,100] window and insignificant in the [101,120] window.

We also conduct robustness test to control for state effects. Specifically, Seasholes and Zhu (2010) point out the issue of geographical return factors in the examination of local advantage. For example, if both sample firms and sample investors cluster in certain areas (e.g., New England or the Bay area), and if stocks of firms in these areas happen to perform well during the sample period, then one can observe a mechanically positive local advantage. To address this concern, we repeat the analysis by controlling for state fixed effects and the results in Table IA1 of the Internet Appendix show that nonlocal disadvantage is both statistically and economically significant in all models.

[25] We also examine the possibility that nonlocal disadvantage is associated with short-term return reversal by repeating the regressions without controlling for lagged returns, or controlling for lagged returns in the previous 1-month window instead of that in the previous 2-week window. The unreported results show that the nonlocal disadvantage remains almost the same.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 4
Panel regressions of stock returns: Alternative measure or sample

	Dependent variable			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
A. Local posts identified using the 50-mile criterion				
Local sentiment	−0.001	−0.006	−0.016	−0.028
	(−0.12)	(−0.37)	(−0.51)	(−0.56)
Nonlocal sentiment	−0.063***	−0.145***	−0.288***	−0.606***
	(−4.91)	(-4.64)	(−5.01)	(−6.34)
Local – Nonlocal	0.062***	0.139***	0.273***	0.578***
	(3.82)	(3.55)	(3.77)	(4.77)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
B. Local posts identified using the state criterion				
Local sentiment	−0.009	−0.023	−0.046	−0.079
	(−0.91)	(−1.05)	(−1.29)	(−1.31)
Nonlocal sentiment	−0.071***	−0.162***	−0.295***	−0.598***
	(−5.12)	(-5.48)	(−5.85)	(−7.60)
Local – Nonlocal	0.062***	0.139***	0.250***	0.519***
	(3.21)	(3.60)	(4.17)	(5.83)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
C. Require at least one post in the sentiment window				
Local sentiment	−0.008	−0.018	−0.058	−0.129
	(−0.52)	(−0.62)	(−1.16)	(−1.53)
Nonlocal sentiment	−0.080***	−0.180***	−0.305***	−0.601***
	(−3.76)	(-4.04)	(−3.78)	(−4.27)
Local – Nonlocal	0.071**	0.161***	0.247**	0.472***
	(2.32)	(2.71)	(2.43)	(2.63)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	238,012	237,942	237,847	237,661
Number of PERMNOs	1,082	1,082	1,082	1,082
D. Include only recommended users				
Local sentiment	−0.014*	−0.038**	−0.076**	−0.178***
	(−1.88)	(−2.10)	(−2.35)	(−2.81)
Nonlocal sentiment	−0.057***	−0.133***	−0.256***	−0.475***
	(−4.84)	(-5.26)	(−5.75)	(−7.48)
Local – Nonlocal	0.042***	0.095***	0.181***	0.297***
	(2.85)	(2.95)	(3.06)	(2.72)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
E. Include only posters with Least Five Local Posts and Five Nonlocal Posts				
Local sentiment	−0.011	−0.028*	−0.062**	−0.120**
	(−1.44)	(−1.69)	(−2.09)	(−2.24)
Nonlocal sentiment	−0.063***	−0.145***	−0.267***	−0.530***
	(−4.76)	(-5.85)	(−4.96)	(−5.61)
Local – Nonlocal	0.051***	0.117***	0.206***	0.410***
	(3.19)	(3.49)	(3.28)	(3.56)

(continued)

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 4
Continued

	Dependent variable			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082

F. Regressions of abnormal returns in longer-term windows

	CAR[21, 40]	CAR[41,60]	CAR[61, 80]	CAR[81,100]	CAR[101,120]
Local sentiment	0.015	−0.037	−0.076*	0.056	0.027
	(0.33)	(−1.50)	(−1.65)	(1.00)	(0.71)
Nonlocal sentiment	−0.592***	−0.493***	−0.290***	−0.170*	−0.068
	(−5.94)	(−4.48)	(−3.31)	(−1.93)	(−0.85)
Local – Nonlocal	0.607***	0.456***	0.214***	0.226*	0.096
	(4.74)	(3.47)	(2.09)	(1.71)	(0.91)
Lagged returns	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Number of obs	621,817	620,801	619,709	618,524	617,242
Number of PERMNOs	1,081	1,081	1,081	1,080	1,080

This table presents the regressions of abnormal returns on local and nonlocal sentiment. Panels A and B are similar to panel A of Table 3, except that we classify Twitter posts into local and nonlocal posts according to whether the Twitter users' locations are within *50 miles* of company headquarters (panel A) or whether the users and the company headquarters locate in the same state (panel B). Panel C presents the regressions of abnormal returns similar to panel A of Table 3, except that we only include firm-days that have at least one Twitter post in the 2-week period of sentiment measurement. Panel D is similar to panel A of Table 3, except that we include only the users recommended by Stocktwits.com. Panel E is similar to panel A of Table 3, except that we only include Twitter posters that post at least five local posts and five nonlocal posts during the sample period. Panel F is similar to panel A of Table 3, except that the independent variables are 20-day abnormal returns in longer windows, from 20 trading days to 120 trading days after construction of the measure. All regressions include firm fixed effects and lagged daily returns of the previous 10 trading days as controls. *t*-statistics (reported in parentheses) are calculated with the Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

2.5 Classifying sentiment in Twitter posts using the Bayesian approach

To corroborate the results using the maximum entropy (ME) approach, we repeat the analysis using a naive Bayesian (NB) approach rather than the ME approach. The existing literature (Li 2010) shows that the NB approach is superior to word count methods for predicting the sentiment of forward-looking statements in corporate filings. Similar to ME estimation, the NB classifier is a maximum likelihood application of Bayes' rule to a set of document features that makes the simplifying assumption that all features are independent. In the field of machine learning, NB is a popular technique to use as a baseline approach for document classification or spam filtering. We follow a similar NB approach to Li (2010) to classify information.[26]

[26] The only difference is that Li (2010) uses four categories (positive, negative, neutral, and uncertain), whereas we maintain the three categories (positive, negative, and neutral) to be consistent with our main tests. Li's uncertain category refers to specific words in his dictionary that referred to uncertainty, rather than being uncertain as to how to classify the document.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

In practice, the NB approach is similar to the ME approach in that it relies on a training set of 2,000 hand-classified posts to determine the probability that each word reflects a positive, negative, or neutral sentiment. Each word is a feature used to classify any document in the full data set. The classification method uses the same equations: (A1) and (A2) in Appendix A.3. Practically, NB is a constrained ME technique; the constraint being that the NB algorithm can only use single words, and not word combinations, to classify sentiment. The robustness test using the NB approach allows us to relate to the classification approach used in the existing literature and provides a strict test of the ME approach used in this paper. Specifically, if the ME approach successfully identifies the information in the sample Twitter posts, then the unconstrained ME approach should perform at least as well as the constrained NB approach.

We repeat the regression analysis in Table 3 but with the investor sentiment based on the NB approach. In panel A of Table 5, the nonlocal disadvantage is robust to the use of NB approach. For example, in the model of 5-day return, the nonlocal disadvantage is 8.5 bps (t-stat = 1.90) per week. As it should be given the constrained nature of the NB approach, this nonlocal disadvantage is smaller than that using the ME approach, but remains both statistically and economically significant. The marked difference in measured nonlocal disadvantage using the NB approach suggests that ME classification may be more effective method of sentiment classification. Panel B of Table 5 presents the tests using the lagged 1-week sentiment measurement window, and the results are similar to panel A. Overall, Table 5 shows that our finding of nonlocal disadvantage is robust to the alternative approach of classifying sentiment.

3. Nonlocal Disadvantage, Information Environment, and Stock Mispricing

The finding of nonlocal disadvantage suggests that nonlocal investors, who suffer more from information asymmetry between firms and investors, make consistent mistakes in predicting stock returns. Therefore, we examine how nonlocal disadvantage varies across alternative information environments.

3.1 News coverage and the nonlocal disadvantage

Mispricing can be stronger among firms with less news coverage. Tetlock (2010), for example, contends that news releases level the playing field for uninformed investors. We examine whether the sentiment errors made by nonlocal investors are affected by news coverage using news articles from the three major news wires: Reuters News, Dow Jones News Wire, and PR News Wire from Factiva. Appendix A.4 provides details about the collection of news stories. We collect 414,301 news articles that cover our 1,082 sample firms during the 2-year sample period. Since our sample comprises relatively large firms, our sample is congruent with that of Fang and Peress (2009) who

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 5

Panel regressions of stock returns on local and nonlocal sentiment: Sentiment of the Twitter posts measured using the Bayesian approach

	Dependent variables			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
A. Regressions of abnormal returns				
Local sentiment	0.001	−0.004	−0.016	−0.028
	(0.14)	(−0.19)	(−0.39)	(−0.45)
Nonlocal sentiment	−0.041***	−0.090***	−0.151***	−0.320***
	(−3.44)	(−3.18)	(−2.88)	(−3.65)
Local – Nonlocal	0.042**	0.085*	0.135	0.292***
	(2.23)	(1.90)	(1.62)	(2.27)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082
B. Regressions of abnormal returns: 1-week sentiment				
Local sentiment	−0.003	−0.002	−0.019	−0.036
	(−0.35)	(−0.12)	(−0.63)	(−0.83)
Nonlocal sentiment	−0.034***	−0.085***	−0.142***	−0.261***
	(−3.76)	(−4.00)	(−3.67)	(−4.05)
Local – Nonlocal	0.032**	0.083***	0.122**	0.225**
	(2.24)	(2.67)	(2.06)	(2.52)
Controls of lagged returns	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
Number of obs	623,748	623,601	623,352	622,825
Number of PERMNOs	1,082	1,082	1,082	1,082

Panel A presents panel regressions of stock returns on prior local and nonlocal sentiment. The dependent variable is cumulative 2-, 5-, 10-, or 20-day abnormal returns (measured in percentage). To calculate daily abnormal return for a firm-day, we first estimate a Fama-French four-factor regression for the firm in the previous 150-day rolling window and then use the estimated factor loadings to calculate abnormal returns for the firm-day. The independent variables include local sentiment and nonlocal sentiment in the 2-week window prior to return measurement. To calculate local and nonlocal sentiment, we first classify Twitter posts into local and nonlocal posts according to whether the Twitter users' locations are within 100 miles of the headquarters of the firms mentioned in the posts. We use the *Bayesian* classification approach to measure the sentiment of each post and then sum the sentiment measures of the local and nonlocal posts, respectively, in the 2 weeks prior to return measurement. We standardize the independent variables for each regression. For each regression, we further report the difference between the coefficients on local sentiment and nonlocal sentiment. Panel B is similar to panel A, except that the local and nonlocal sentiment is measured in the 1-week window prior to return measurement. All regressions include firm fixed effects with lagged daily returns in the previous 10 trading days as controls. *t*-statistics (reported in parentheses) are calculated using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

examine a sample of large firms (NYSE stocks plus 500 randomly selected NASDAQ stocks). The annual news coverage rate for our sample is 98.0%, higher than that in Fang and Peress (2009), who find that annual news coverage by four nationwide newspapers ranges between 57% and 77% during 1993 to 2002.[27]

[27] We attribute our higher media coverage rate to three factors. Our sample period is almost a decade later than that of Fang and Peress (2000); our news wires represent a broader set of publications than the four newspapers used by Fang and Peress (2009); and because Twitter activity is related to news (panel C of Table 1 shows a strong positive relation between news coverage and posting activity), this fact could reinforce our high coverage rate.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

For each day of our sample period, we divide firms into two groups based on whether the firms have news coverage in the previous 2 weeks (the measurement window for Twitter sentiment). Panel A of Table 6 presents the regressions of abnormal stock returns for the no-news and news subsamples. We observe that nonlocal disadvantage is much larger in the no-news sample than in the news sample. For example, nonlocal disadvantage is 45.4 bps per week (t-stat = 4.57) for the no-news sample, but only 10.2 bps (t-stat = 3.03) for the news sample. The spread of nonlocal disadvantage is large and significant at 35.2 bps (t-stat = 3.36). More importantly, the larger nonlocal disadvantage in the no-news sample is driven by the much stronger contrarian predictive ability of nonlocal sentiment. This finding suggests that news coverage helps to alleviate nonlocal disadvantage.

3.2 Information asymmetry and the nonlocal disadvantage

Next, we examine how nonlocal disadvantage varies with the level of information asymmetry between the firm and investors. In particular, we examine three commonly used proxies for information asymmetry proposed in previous studies beginning with firm size since Coval and Moskowitz (1999) and Hong, Lim, and Stein (2000) contend that small firms have greater information asymmetry than large firms. For each day in our sample period, we classify firms into quartiles according to their market capitalizations at the end of the previous year and calculate nonlocal disadvantage for small (bottom quartile) and large firms (top quartile) separately. Panel B of Table 6 shows that nonlocal disadvantage for small firms is significantly larger than that of the large firms, and that it is largely driven by the contrarian predictive ability of nonlocal sentiment.

Our second proxy for information asymmetry is analyst coverage as existing studies suggest that firms followed by larger numbers of analysts tend to have lower information asymmetry (e.g., Brennan and Subrahmanyam 1995; Hong, Lim, and Stein 2000; Irvine 2004). Since analyst coverage and firm size are strongly correlated, we construct size-adjusted analyst coverage as the residual from a cross-sectional regression of analyst coverage on firm size. For each day in our sample period, we sort firms into quartiles according to their size-adjusted analyst coverage for the month, and examine nonlocal disadvantage for low coverage firms (the bottom quartile of coverage) and high coverage firms (the top quartile of coverage). Panel C of Table 6 shows that nonlocal disadvantage for low-coverage firms is significant in all return windows and much larger than that of the full sample (panel A of Table 3). In contrast, nonlocal disadvantage is much smaller for the high coverage firms. The difference is again driven by the contrarian predictive ability of nonlocal sentiment in low coverage stocks.

Our third proxy for information asymmetry is S&P 500 index membership, as stocks in the index are generally associated with more media exposure and

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 6

Panel regressions of stock returns: Subsample analysis across public news or information asymmetry proxies

	Dependent variables			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
A. Nonlocal disadvantage for no-news versus news Firms				
Firms with no news				
Local sentiment	0.016	−0.016	−0.160*	−0.419***
	(0.49)	(−0.23)	(−1.66)	(−2.90)
Nonlocal sentiment	−0.187***	−0.470***	−0.891***	−1.548***
	(−5.90)	(−5.96)	(−5.24)	(−5.26)
Local - Nonlocal (1)	0.203***	0.454***	0.731***	1.129***
	(4.90)	(4.57)	(3.85)	(3.41)
Firms with news				
Local sentiment	−0.010	−0.019	−0.035	−0.057
	(−1.19)	(−1.21)	(−1.22)	(−1.07)
Nonlocal sentiment	−0.054***	−0.121***	−0.219***	−0.460***
	(−4.29)	(−4.58)	(−4.41)	(−5.45)
Local - Nonlocal (2)	0.044***	0.102***	0.184***	0.402***
	(2.54)	(3.03)	(2.92)	(3.64)
No-news − News				
(1) - (2)	0.159***	0.352***	0.547***	0.727**
	(3.53)	(3.36)	(2.73)	(2.08)
B. Nonlocal disadvantage for small versus large firms				
Small firms				
Local sentiment	−0.056	−0.232**	−0.555***	−1.231***
	(−1.47)	(−2.58)	(−3.40)	(−3.37)
Nonlocal sentiment	−0.387***	−0.879***	−1.582***	−2.987***
	(−5.80)	(−5.63)	(−5.24)	(−5.53)
Local - Nonlocal (1)	0.332***	0.647***	1.027***	1.756**
	(3.94)	(3.33)	(2.81)	(2.56)
Large firms				
Local sentiment	0.006	0.017	0.021	0.038
	(0.78)	(1.15)	(0.86)	(0.78)
Nonlocal sentiment	−0.029***	−0.066***	−0.114**	−0.262***
	(−2.56)	(−2.60)	(−2.41)	(−2.87)
Local - Nonlocal (2)	0.035**	0.083**	0.134**	0.299**
	(2.05)	(2.33)	(2.12)	(2.36)
Small − Large				
(1) - (2)	0.297***	0.564***	0.893**	1.456**
	(3.46)	(2.86)	(2.43)	(2.09)
C. Nonlocal disadvantage for low versus high analyst coverage firms				
Low coverage firms				
Local sentiment	−0.055**	−0.149***	−0.317***	−0.636***
	(−2.02)	(−2.75)	(−3.79)	(−4.09)
Nonlocal sentiment	−0.197***	−0.498***	−0.897***	−1.771***
	(−3.83)	(−4.80)	(−5.16)	(−6.24)
Local - Nonlocal (1)	0.142***	0.350***	0.580***	1.136***
	(2.18)	(2.74)	(2.70)	(3.23)
High coverage firms				
Local sentiment	−0.003	0.005	0.005	0.019
	(−0.30)	(0.20)	(0.13)	(0.29)
Nonlocal sentiment	−0.032**	−0.074**	−0.120*	−0.259**
	(−2.22)	(−2.18)	(−1.90)	(−2.36)
Local - Nonlocal (2)	0.029	0.079*	0.125	0.278*
	(1.37)	(1.67)	(1.49)	(1.88)
Low − High				
(1) - (2)	0.113*	0.271**	0.455**	0.858***
	(1.65)	(1.99)	(1.97)	(2.25)

(continued)

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 6
Continued

	Dependent variables			
	2-day CAR	5-day CAR	10-day CAR	20-day CAR
D. Nonlocal disadvantage for S&P 500 index firms versus non-index firms				
Non-index firms				
Local sentiment	−0.038**	−0.095**	−0.195***	−0.383***
	(−2.19)	(−2.57)	(−3.01)	(−2.89)
Nonlocal sentiment	−0.118***	−0.271***	−0.496***	−0.940***
	(−4.97)	(−5.08)	(−4.94)	(−5.17)
Local - Nonlocal (1)	0.080**	0.175**	0.300**	0.557**
	(2.52)	(2.56)	(2.29)	(2.29)
Index firms				
Local sentiment	0.006	0.015	0.021	0.047
	(0.80)	(1.07)	(0.86)	(1.07)
Nonlocal sentiment	−0.032***	−0.075**	−0.139**	−0.322**
	(−2.67)	(−2.67)	(−2.57)	(−3.22)
Local - Nonlocal (2)	0.038**	0.090**	0.160**	0.370
	(2.29)	(2.54)	(2.40)	(2.95)
Non-index − Index				
(1) - (2)	0.041	0.086	0.140	0.187
	(1.16)	(1.11)	(0.96)	(0.69)

Panel A separately reports regressions of 2-, 5-, 10-, or 20-day abnormal returns on local and nonlocal sentiment in the 2-week period prior to return measurement for stocks with and without public news coverage. We collect news articles from the PR News Wire, Dow Jones News Wire, and Reuters News and classify stocks into two groups based whether they have news coverage in the 2-week period of sentiment measurement. We then estimate regressions for the no-news firms and the news firms separately. We further report the difference in nonlocal disadvantage between no-news and news samples. The regression settings and the independent variables are defined in the heading of Table 3. Panels B reports regressions for small and large firms. On each day of our sample period, we sort stocks into four groups based on their market capitalizations. We then estimate regressions of abnormal returns like in the panel A of Table 3 for small firms (lowest quartile of market capitalization) and large firms (highest quartile of market capitalization), respectively. For panel C, on each day of our sample period, we sort stocks into four groups based on size-adjusted analyst coverage, where size-adjusted analyst coverage is residual from cross-sectional regression of analyst coverage on size. We then report nonlocal disadvantage for low coverage firms (lowest quartile of coverage), high coverage firms (highest quartile of coverage), and their differences. For panel D, we classify stocks into two groups based on whether or not they are in the S&P 500 index. We then report nonlocal disadvantage for nonindex firms, index firms, and their differences. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

higher investor interest and therefore lower information asymmetry (Seasholes and Zhu 2010). We sort firms into two groups according to whether they are in the S&P 500 index, and examine nonlocal disadvantage among the two groups separately. Panel D of Table 6 shows that nonlocal disadvantage among non-index firms is generally twice as large as that among index firms, although the difference is not statistically significant. Overall, these results are consistent with information asymmetry exacerbating the nonlocal disadvantage.[28]

[28] We also repeat the regression analysis in the samples double sorted by whether there is news in the sentiment measure period and the proxies of information asymmetry. The difference of nonlocal disadvantage in no news versus news is generally of similar magnitude across both high and low information asymmetry firms, with the only exception of analyst coverage. This evidence suggests that the initial sort on public news articles makes more of a difference in returns than does the secondary sort using asymmetric information proxies. Therefore, the existence of news is the most important proxy for anchoring nonlocal sentiment to a more realistic appraisal of value. For brevity, we report the results in Internet Appendix Table IA2.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

3.3 The relation between the nonlocal disadvantage and mispricing

In the case of nonlocal disadvantage, the different return predictive ability between locals and nonlocals is attributable to stock mispricing. Specifically, the negative return predictive ability of nonlocals can be caused by nonlocals being too optimistic (pessimistic) about overpriced (underpriced) stocks. We conduct two analyses to examine whether nonlocal disadvantage is indeed associated with stock mispricing.

Our first analysis is based on the literature suggesting that mispricing is associated with limits to arbitrage, where sophisticated investors are reluctant to trade against mispricing. A recent work by Stambaugh, Yu, and Yuan (2015) suggests that mispricing, especially overpricing, is stronger in stocks with higher idiosyncratic stock return volatilities. This is because return volatility deters arbitrage ("limits to arbitrage") and arbitragers are more likely to buy underpriced stocks than short overpriced stocks ("arbitrage asymmetry"). We therefore examine the relation between nonlocal disadvantage and idiosyncratic return volatility.

We first estimate idiosyncratic return volatility for a firm-day as standard deviation of the residuals from the time-series regression of daily stock returns on the market factor (MKT) in the one-year window up to the end of previous month. We then sort firms into quartiles based on idiosyncratic return volatility, and examine nonlocal disadvantage among high volatility firms (the top quartile of volatility) and low volatility firms (the bottom quartile of volatility), respectively. Table 7 shows that nonlocal disadvantage among high volatility firms is significant over all return windows examined. On the contrary, nonlocal disadvantage among low volatility firms is small and statistically insignificant. Consistent with the idea that the predictive ability of nonlocal sentiment identifies overpriced stocks, the mispricing identified by nonlocal sentiment appears to be arbitraged away when idiosyncratic volatility is low.[29]

Our second analysis of mispricing examines the relation between Twitter sentiment and subsequent earnings announcement return. This test is based on La Porta et al. (1997), who suggest that if an anomaly is associated with mispricing, then it will be stronger in the earnings announcement window, as the release of earnings helps to correct mispricing.[30] We calculate earnings announcement return as the cumulative abnormal return in the 2-day window [0,1] where day 0 is earnings announcement day.

[29] We also consider the setting in Stambaugh, Yu, and Yuan (2012), who suggest that mispricing, especially overpricing is increasing in market wide sentiment, and they find that anomalies are stronger in the periods of high market sentiment, measured by the Baker-Wurgler (BW) market sentiment index. The test using market-wide sentiment is challenging in our setting, as our sample period is less than three years, and for most of this period the BW sentiment measures fall in the bottom quartile of the sentiment series. We nevertheless conduct this analysis, and the unreported result shows no obvious relation between BW sentiment and nonlocal disadvantage in our sample period.

[30] A contemporaneous study by Engelberg, McLean, and Pontiff (2016) uses this approach to study a strategy that combines 94 anomalies.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 7
Panel regressions of stock returns: Subsample analysis across idiosyncratic return volatility

	Dependent Variables			
	2-Day CAR	5-Day CAR	10-Day CAR	20-Day CAR
High volatility Firms				
Local sentiment	−0.035*	−0.073**	−0.146**	−0.302**
	(−1.87)	(−2.01)	(−2.10)	(−2.13)
Nonlocal sentiment	−0.158***	−0.318***	−0.546***	−1.018***
	(−4.47)	(−4.13)	(−3.70)	(−3.81)
Local - Nonlocal (1)	0.124***	0.235***	0.400**	0.716**
	(3.06)	(2.67)	(2.34)	(2.27)
Low volatility firms				
Local sentiment	0.000	−0.001	−0.013	−0.022
	(0.08)	(−0.08)	(−0.64)	(−0.62)
Nonlocal sentiment	−0.007	−0.015	−0.014	−0.060
	(−0.92)	(−0.85)	(−0.42)	(−1.06)
Local - Nonlocal (2)	0.008	0.014	0.000	0.037
	(0.65)	(0.54)	(0.01)	(0.49)
High vol. – Low vol.				
(1) - (2)	0.116***	0.221***	0.400**	0.679**
	(2.75)	(2.41)	(2.26)	(2.10)

This table reports regressions of 2-, 5-, 10-, or 20-day abnormal returns on local and nonlocal sentiment in the 2-week period prior to return measurement for stocks with high return volatilities and stocks with low return volatilities separately. On each day of our sample period, we sort stocks into four groups based on idiosyncratic volatility. Idiosyncratic volatility for a firm-day is standard deviation of the residuals from the time-series regression of daily stock returns on the market factor (MKT) in the 1-year window up to the end of previous month. We then estimate regressions for high volatility firms (highest quartile of volatility) and low volatility firms (lowest quartile of volatility) separately. We further report the difference in nonlocal disadvantage between high-volatility and low-volatility samples. The regression setting and the variables are as defined in the heading of Table 3. ***, **, and * represent statistical significances at the 0.01, 0.05, and 0.10 levels, respectively.

The left column of Table 8 presents a cross-sectional regression of earnings announcement returns on local and nonlocal sentiment in the previous 2-week window for 9,735 earnings announcements with available data for our sample firms. The local and nonlocal sentiment is standardized to facilitate the comparison of economic significance. The coefficient on local sentiment is insignificant, whereas that on nonlocal sentiment is significantly negative. More importantly, the nonlocal disadvantage in terms of 2-day earnings announcement return is 0.478% (t-stat = 2.78), about 7.5 times as large as the nonlocal disadvantage of 0.064% for an ordinary 2-day return (panel A of Table 3). The right column repeats the regression analysis using 1-week local and nonlocal sentiment as independent variables, and the results are similar. Therefore, the results based on idiosyncratic volatility and earnings announcement return are both consistent with the mispricing explanation for nonlocal disadvantage.

4. Nonlocal Disadvantage and Stock Overpricing

4.1 Long-short trading strategies based on the nonlocal disadvantage

Our evidence so far suggests that nonlocal disadvantage is associated with stock mispricing. A number of studies have documented that because of short

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 8
Regressions of earnings announcement returns on investor sentiment

	Dependent variable: CAR [0,1] (%)	
	2-week sentiment	1-week sentiment
Local sentiment	0.043	−0.013
	(0.38)	(−0.11)
Nonlocal sentiment	−0.435***	−0.394***
	(−3.43)	(−3.15)
Local - Nonlocal	0.478***	0.381**
	(2.78)	(2.20)
Firm fixed effects	Yes	Yes
Number of obs	9,735	9,735
Number of PERMNOs	1,068	1,068
R^2	0.125	0.125

This table presents regression of earnings announcement returns (EARs) on local and nonlocal sentiment. The sample includes 9,735 quarterly earnings announcements of sample firms that have available data to estimate unexpected earnings. The dependent variable is cumulative abnormal return (CAR) in the 2-day window [0,1] around earnings announcement, where day 0 is the earnings announcement day. For the left (right) column, the independent variables include local and nonlocal sentiment measured in the two weeks (one week) prior to the earnings announcement date. To facilitate the comparison of economic significance, we standardize the local and nonlocal sentiment. The regressions also include firm fixed effects. *t*-statistics are reported in parentheses. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

sales constraints, overpricing dominates stock mispricing (e.g., Miller 1977; Stambaugh, Yu, and Yuan 2015). We therefore examine the profitability of zero-investment trading strategies based on nonlocal disadvantage. This examination is not only of interest to practitioners but also helps us investigate whether the nonlocal disadvantage comes from the short side or the long side.

Before examining nonlocal advantage, we first examine a zero-investment strategy based on total Twitter sentiment. Since the total Twitter sentiment negatively predicts stock return, this strategy goes long the unconditionally unfavorable portfolio and goes short the unconditionally favorable portfolio. Specifically, for each firm-day, we calculate total investor sentiment in the previous 2 weeks and form a portfolio containing firms for which the sentiment measures are greater than zero ("unconditionally favorable"), and a portfolio containing firms for which the sentiment measures are less than or equal to zero ("unconditionally unfavorable"). We then hold these portfolios for J days, where $J = 2$, 5, 10, or 20. This strategy is similar to the rolling momentum strategy proposed by Jegadeesh and Titman (1993), except that we form portfolios based on Twitter sentiment rather than momentum. Daily abnormal returns are constructed based on the four-factor model defined in Section 2. We then calculate the daily abnormal profit of the long-short strategy.

Table 9 shows that the daily abnormal profit of the long-short strategy ranges from 6.3 bps to 6.8 bps and is statistically significant for all holding windows. Furthermore, the profit is driven by the negative return of short portfolio, suggesting that the negative return predictive ability of Twitter sentiment is due to posters favoring overpriced stocks.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 9

Daily abnormal profit (%) of rolling long-short strategy based on unconditionally favorable versus unfavorable investor sentiment

Hold 2 days			Hold 5 days			Hold 10 days			Hold 20 days		
Short	Long	L-S	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S
−0.071***	−0.004	0.067***	−0.071***	−0.004	0.068***	−0.074***	−0.007	0.067***	−0.070***	−0.008	0.063***
(−3.89)	(−0.47)	(4.04)	(−3.88)	(−0.44)	(4.14)	(−4.13)	(−0.80)	(4.34)	(−4.20)	(−0.86)	(4.52)

This table presents daily abnormal profit (%) of rolling long-short strategy based on investor sentiment. For each firm-day, we calculate investor sentiment in the previous 2 weeks and form a portfolio containing firms for which the sentiment measures are greater than zero ("unconditionally favorable," short), as well as a portfolio containing firms for which the sentiment measures are less than or equal to zero ("unconditionally unfavorable," long). We then hold these portfolios for J days, where $J = 2$, 5, 10, or 20. This strategy is similar to the rolling momentum strategy proposed by Jegadeesh and Titman (1993), except that we form portfolios based on differential sentiment rather than momentum. We then calculate the daily abnormal profits of a strategy that long the "unconditionally unfavorable portfolio" and short the "unconditionally favorable portfolio." Specifically, we first calculate for each day the difference in average abnormal returns between the two portfolios and then report time-series means of the daily abnormal profits. Daily abnormal return is constructed based on Fama and French four-factor model and is defined in the heading of Table 2. We report daily abnormal profit of the long-short strategy. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Next, we examine a long-short strategy based on nonlocal disadvantage, where we form two portfolios based on the contrast between nonlocal sentiment and local sentiment. The first portfolio, "nonlocal favorable portfolio," contains firms for which the difference between nonlocal and local sentiment in the past 2 weeks is greater than zero. The second portfolio, "nonlocal unfavorable portfolio," contains firms for which the difference between nonlocal and local sentiment in the past 2 weeks is less than or equal to zero. We then hold the two portfolios for J days, where $J = 2$, 5, 10, or 20.

Table 10 reports the average daily abnormal profit of a zero-investment strategy that goes long the nonlocal unfavorable portfolios and goes short the nonlocal favorable portfolios. The daily abnormal profit ranges from 5.4 bps to 6.1 bps and is statistically significant for all windows. Additionally, consistent with the overpricing explanation, the profit is driven by the short portfolio.

Our previous results show that nonlocal disadvantage varies across firm characteristics including firm size, analyst coverage, index membership, and idiosyncratic volatility. To illustrate the potential magnitude of the nonlocal-disadvantage-based strategy, we further exclude large firms, high analyst coverage firms, low idiosyncratic volatility firms, or index firms (their classifications are defined in Section 3). Table 10 shows that the daily abnormal profits increase in these subsamples. For example, excluding the large firms (the top quartile of firm size), the daily abnormal profit ranges from 7.9 bps to 8.6 bps across different holding windows.

A natural question is whether the profit of long-short strategy exceeds the transaction cost. The cost of short selling is particularly relevant given that the bulk of the profit from the strategy comes from the short side of the portfolio. Our results are consistent with Miller (1977) in that given the existence of short sale constraints, optimists can cause overpricing of securities. Our results are also consistent with the anomaly results reported in Stambaugh, Yu, and Yuan (2012), where most of the abnormal profits come from the short-side of the anomaly portfolios. Reed (2013) echoes D'Avolio (2002) and notes that the primary cost of short selling is the discount the short seller receives on the funds they must deposit with the broker relative to the fed funds rate. Reed (2013) reports the typical discount is between 5-25 points below the federal funds rate on an annual basis. Such a modest cost should not significantly affect the profit of our reported long-short strategy over the horizons we examine. D'Avolio (2002) reports that 91% of stocks typically can be borrowed for less than 1% per annum but significantly higher discounts can occur when the supply of lendable shares is low. He finds that high borrowing costs occur primarily in smaller stocks. These smaller and potentially more costly to short stocks are not an important

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 10
Daily abnormal profit (%) of rolling long-short strategy based on the nonlocal disadvantage

Hold 2 days			Hold 5 days			Hold 10 days			Hold 20 days		
Short	Long	L-S	Short	Long	L-S	Short	Long	L-S	Short	Long	L-S
A. Full sample											
−0.067***	−0.007	0.061***	−0.067***	−0.007	0.061***	−0.069***	−0.008	0.060***	−0.064***	−0.011	0.054***
(−3.86)	(−0.28)	(3.76)	(−3.75)	(−0.77)	(3.80)	(−3.88)	(−1.13)	(3.83)	(−3.99)	(−1.21)	(3.82)
B. Exclude large firms											
−0.088***	−0.005	0.084***	−0.089***	−0.004	0.086***	−0.090***	−0.007	0.084***	−0.085***	−0.007	0.079***
(−3.17)	(−0.45)	(3.21)	(−3.43)	(−0.40)	(3.67)	(−3.50)	(−0.63)	(3.71)	(−3.53)	(−0.71)	(3.83)
C. Exclude high analyst coverage firms											
−0.083***	−0.008	0.076***	−0.097***	−0.007	0.090***	−0.102***	−0.010	0.092***	−0.097***	−0.011	0.087***
(−3.68)	(−0.87)	(3.57)	(−4.66)	(−0.83)	(4.62)	(−4.87)	(−1.14)	(4.80)	(−5.10)	(−1.15)	(5.11)
D. Exclude S&P 500 index firms											
−0.091***	−0.009	0.084***	−0.090***	−0.008	0.083***	−0.098***	−0.011	0.088***	−0.099***	−0.011	0.089***
(−3.09)	(−0.78)	(3.05)	(−3.31)	(−0.77)	(3.37)	(−3.76)	(−1.02)	(3.83)	(−4.02)	(−1.07)	(4.17)
E. Exclude low idiosyncratic volatility firms											
−0.080***	−0.006	0.076***	−0.082***	−0.006	0.077***	−0.085***	−0.009	0.077***	−0.086***	−0.011	0.076***
(−3.37)	(−0.49)	(3.54)	(−3.38)	(−0.51)	(3.62)	(−3.55)	(−0.84)	(3.70)	(−3.90)	(−0.95)	(4.15)

Panel A presents daily abnormal profit (%) of rolling long-short strategy based on local versus nonlocal sentiment. For each firm-day, we contrast the nonlocal sentiment and local sentiment in the previous two weeks. Then on each day, we form a portfolio containing firms for which nonlocal sentiment is higher than local sentiment ("nonlocal favorable portfolio," short), as well as a portfolio containing firms for which the nonlocal sentiment is less than or equal to local sentiment ("nonlocal unfavorable portfolio," long). We then hold these portfolios for J days, where $J = 2, 5, 10$, or 20. This strategy is similar to the rolling momentum strategy proposed by Jegadeesh and Titman (1993), except that we form portfolios based on differential sentiment rather than momentum. We then calculate the daily abnormal profit of a strategy that long the "nonlocal unfavorable portfolio" and short the "nonlocal favorable portfolio." Specifically, we first calculate for each day the difference in average abnormal returns between the two portfolios, and then report time-series means of the daily abnormal profits. Daily abnormal return is constructed based on Fama and French four-factor model and is defined in the heading of Table 2. Panels B to E report daily abnormal profits of the long-short strategy for the subsamples that exclude large firms, high analyst coverage firms, S&P 500 index firms, or low idiosyncratic volatility firms. The classifications of subsamples are defined in the heading of Tables 6 and 7. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

component of our strategies given that our posters have a strong tendency to post only about large capitalization stocks (see Table 1).[31]

Regarding execution costs, Hatheway, Kwan, and Zheng (forthcoming) examine a sample of large stocks in the first quarter of 2011 (a period that coincides with our study), and report effective spreads ranging from 1.13 to 1.39 bps per trade. These execution costs indicate that our local-nonlocal investment strategy would be marginally profitable if executed over a 2-day holding period. As we find that profit from the strategy continues to increase for up to 20 trading days, the profit of the strategy can clearly cover the transaction cost of the trade.[32]

4.2 Relation of Twitter sentiment to contemporaneous trading volume and stock returns

Why is nonlocal sentiment higher in overpriced stocks? It is possible that nonlocals, who lack sophistication and information, focus on "glamor" stocks, which tend to be overpriced. To illustrate this possibility, we examine whether local sentiment and nonlocal sentiment respond to contemporaneous trading volume and stock return differently.

We first examine the contemporaneous relation between Twitter sentiment and trading volume. In panel A of Table 11, we estimate regressions of average daily abnormal trading volume in the 2-week sentiment measurement window. The dependent variable is average daily abnormal turnover. For a firm-day, we first calculate daily turnover as a firm's daily trading volume scaled by total shares outstanding, and then calculate daily excess turnover by subtracting cross-sectional average turnover of the CRSP universe. We then calculate abnormal turnover for a firm-day by subtracting average daily excess turnover of the firm over the previous 180 days. The independent variables include local sentiment and nonlocal sentiment in the 2-week window as defined in previous sections.

The results in panel A show that although both local sentiment and nonlocal sentiment are positively associated with contemporaneous trading volume, nonlocal sentiment is much more strongly related to volume than local sentiment. In panel B, we turn to the daily level and regress daily abnormal turnover on the same-day local and nonlocal sentiment, and the results also show that nonlocal sentiment has a much stronger relation with contemporaneous volume than does local sentiment.

We further examine the contemporaneous relation between Twitter sentiment and stock return in Table 12. In panel A, we estimate regressions of

[31] It is notable that D'Avolio (2002) includes a measure of Yahoo! message board activity and finds that poster activity has a weak positive relation to shorting costs. However, the relation is not monotonic and only has an economically significant effect for small stocks.

[32] Additionally, Table 10 shows that stocks likely to be more expensive to short are also associated with higher profits (see D'Avolio (2002) who documents the positive relation of short sale costs with size and idiosyncratic volatility). Therefore, the exact "sweet spot" for the trading strategy is indeterminate.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 11

Panel regressions of contemporaneous trading volume on local and nonlocal sentiment

A. Regressions of average daily abnormal turnover (%) on local and nonlocal sentiment

	Dep. var.: Average daily CAT in the sentiment window		
Local sentiment	0.067***		0.022
	(5.07)		(1.52)
Nonlocal sentiment		0.160***	0.153***
		(7.48)	(6.49)
Nonlocal – Local			0.131***
			(4.06)
Controls of lagged returns	Yes	Yes	Yes
Firm Effects	Yes	Yes	Yes
Number of obs	623,843	623,843	623,843
Number of PERMNOs	1,082	1,082	1,082

B. Regressions of daily abnormal turnover (%) on the same-day sentiment

	Dep. var.: Daily CAT		
Daily local sentiment	0.075***		0.043***
	(5.95)		(3.59)
Daily nonlocal sentiment		0.237***	0.230***
		(9.35)	(9.02)
Nonlocal - Local			0.187***
			(6.57)
Controls of lagged returns	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes
Number of obs	623,793	623,793	623,793
Number of PERMNOs	1,082	1,082	1,082

Panel A presents panel regressions of average daily abnormal trading volume in the 2-week measurement window of the local and nonlocal sentiment. The dependent variable is average daily abnormal turnover (measured in percentage). Daily turnover is a firm's daily trading volume scaled by total shares outstanding. We first obtain daily excess turnover by subtracting cross-sectional average turnover of the CRSP universe and then calculate abnormal turnover for a firm-day by subtracting average daily excess turnover of the firm in the previous 180-day rolling window. The independent variables include local sentiment and nonlocal sentiment measures in the two-week window defined in the heading of Table 3. The regression setting is the same as panel A of Table 3, except the dependent variable is abnormal turnover. Panel B is similar to panel A, except that the dependent variable is the daily abnormal turnover, and the independent variables include the same-day local and nonlocal sentiment. All regressions include firm fixed effects with lagged returns in the previous ten trading days as controls. *t*-statistics (reported in parentheses) are calculated using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

average daily abnormal return in the 2-week measurement window of the local and nonlocal sentiment. In the full regression that includes both local and nonlocal sentiment, the coefficient on local sentiment is insignificant, whereas that on nonlocal sentiment is significantly positive. The difference between local and nonlocal is also statistically significant. Since Stambaugh, Yu, and Yuan (2015) show that overpricing is more prevalent in high volatility stocks, we repeat the regression for high and low volatility stocks, respectively. Panel A shows that the results are much stronger in high volatility stocks where overpricing is strong. Panel B repeats the analysis using daily abnormal return and same-day sentiment, and the results are similar.

Overall, the results in this subsection shows that compared to local posters, nonlocal poster sentiment seems to focus on "glamor" stocks, which can

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 12

Panel regressions of contemporaneous returns on local and nonlocal sentiment

A. Regressions of average daily AR (%) on local and nonlocal sentiment

	Dep. var.: Average daily AR in the sentiment window				
	Full sample			Low volatility	High volatility
Daily local sentiment	0.014***		0.006	−0.002	0.006
	(3.05)		(1.20)	(-0.58)	(0.45)
Daily nonlocal sentiment		0.030***	0.028***	0.014***	0.079***
		(4.08)	(3.49)	(2.97)	(4.28)
Nonlocal - Local			0.022**	0.016**	0.073***
			(2.14)	(2.29)	(3.29)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Number of obs	624,057	624,057	624,057	155,501	155,761
Number of PERMNOs	1,082	1,082	1,082	421	533

B. Regressions of daily AR (%) on the same-day sentiment

	Dep. var.: Daily AR				
	Full sample			Low volatility	High volatility
Daily local sentiment	0.053***		0.032***	−0.003	0.074**
	(4.94)		(2.78)	(-0.48)	(2.33)
Daily nonlocal sentiment		0.159***	0.154***	0.053***	0.270***
		(12.43)	(11.82)	(5.68)	(7.37)
Nonlocal - Local			0.122***	0.056***	0.196***
			(6.58)	(5.02)	(4.27)
Lagged returns	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
Number of obs	623,797	623,797	623,797	155,499	155,704
Number of PERMNOs	1,082	1,082	1,082	421	533

Panel A presents panel regressions of average daily abnormal return in the 2-week measurement window of the local and nonlocal sentiment. The dependent variable is average daily abnormal return (measured in percentage). Daily abnormal return is defined in the heading of Table 2. The independent variables include local sentiment and nonlocal sentiment measures in the 2-week window defined in the heading of Table 3. The regression setting is the same as panel A of Table 3, except the dependent variable is contemporaneous return in the sentiment measurement window. Panel B is similar to panel A, except that the dependent variable is the daily abnormal return, and the independent variables include the same-day local and nonlocal sentiment. All regressions include firm fixed effects with lagged returns in the previous ten trading days as controls. t-statistics (reported in parentheses) are calculated using Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

potentially explain the positive association between nonlocal sentiment and overpricing.

4.3 Nonlocal sentiment and analyst opinions

The previous subsection focuses on the relation between Twitter sentiment and stock market metrics, but this subsection examines whether local and nonlocal sentiment also differently respond to analyst opinions. Specifically, we exploit the existing literature that retail investors fail to correct for the complexity of analysts' incentives and tend to slavishly trade in the direction of the recommendation. For example, Malmendier and Shanthikumar (2007, 2014) and Mikhail, Walther, and Willis (2007) sort investors by trade size and

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

find that small traders tend to overreact to analysts' recommendations, especially buy recommendations, and therefore experience significant underperformance compared to large traders. We therefore examine the responses of local and nonlocal investors to overoptimistic analyst forecasts and analyst recommendations.

We follow the literature (Bradshaw, Richardson, and Sloan 2006) and construct monthly measure of analyst optimism as mean earnings forecast minus the corresponding actual earnings, scaled by stock price at the summary date. Both the mean forecast and actual earnings are obtained from the IBES monthly summary file.[33] We then examine local and nonlocal investors' responses to analyst optimism in a regression setting.

Panel A of Table 13 reports panel regressions of monthly local or nonlocal sentiment on analyst optimism in the previous month for 1,049 firms with available data in the sample period. The independent variable is the sum of sentiment of all Twitter posts from local or nonlocal investors for a firm-month. We standardize both the dependent and the independent variables to facilitate the comparison of economic significance. We observe that the coefficient on the local sentiment is an insignificant 0.008 (t-stat = 0.83), but the coefficient on the nonlocal sentiment is a significantly positive 0.043 (t-stat = 6.96). The difference between the two coefficients is statistically significant at the 0.01 level. These results suggest that nonlocal investors respond much more aggressively to overoptimistic analyst forecasts than do local investors.

We also examine whether nonlocal investors react more strongly to analysts' recommendations than local investors do. We first obtain monthly consensus recommendations from the IBES summary file, where an individual analyst recommendation takes the value of 1 (strong buy), 2 (buy), 3 (hold), 4 (sell), or 5 (strong sell). We then construct a binary variable "sell recommendation" ("buy recommendation") that equals 1 if the consensus recommendation is higher (lower) than 3, and 0 otherwise.

Panel B of Table 13 presents panel regressions of monthly local or nonlocal sentiment on lagged monthly buy and sell recommendations. We also standardize the sentiment variables to facilitate the comparison of coefficients. The coefficients on sell recommendations is insignificantly negative for local sentiment (−0.062, t-stat = −1.04), but significantly negative for nonlocal sentiment (−0.074, t-stat = −2.43). This result indicates that both locals and nonlocals tweet negatively about firms with consensus sell recommendations. Although the coefficient for nonlocal investors is slightly larger, the difference is not statistically significant. For buy recommendations, the coefficient is significantly positive for both local sentiment (0.039, t-stat = 3.39) and nonlocal sentiment (0.095, t-stat = 6.07). The significant difference between the coefficients (0.057, t-stat = 2.91) indicates that nonlocal

[33] We further adjust analyst optimism by controlling for the average of other firms in the same two-digit SIC industry to control for any industry effect.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Table 13

Panel regressions of monthly local or nonlocal sentiment on analyst optimism and consensus analyst recommendation

	Dependent variables		
	Local sentiment	Nonlocal sentiment	Nonlocal - Local
A. Regressions on analyst optimism			
Analyst optimism	0.008	0.043***	0.034***
	(0.83)	(6.96)	(2.94)
Firm fixed effects	Yes	Yes	
Number of obs	26,936	26,936	
Number of PERMNOs	1,049	1,049	
B. Regressions on buy and sell recommendations			
Sell recommendation	−0.062	−0.074**	0.012
	(−1.04)	(−2.43)	(0.17)
Buy recommendation	0.039***	0.095***	0.057***
	(3.39)	(6.07)	(2.91)
Firm fixed effects	Yes	Yes	
Number of obs	30,354	30,354	
Number of PERMNOs	1,063	1,063	

Panel A reports panel regressions of monthly local or nonlocal sentiment on lagged analyst optimism measure. The dependent variable is local or nonlocal sentiment for each firm-month in the sample period. The independent variable is the analyst optimism measure in the month prior to the month of return, where the analyst optimism measure is calculated as mean analyst forecast (obtained from the I/B/E/S summary file) minus the corresponding actual earnings, scaled by stock price of the summary date. We further adjust analyst optimism of a firm by the average of other firms in the same two-digit SIC industry. We standardize the independent and dependent variables to facilitate the comparison of economic significances, and include firm fixed effects in both regressions. *t*-statistics (reported in parentheses) are calculated with Driscoll and Kraay (1998) robust standard errors that control for both cross-sectional and time-series correlations. We also report the difference in the coefficients on analyst optimism and the associated *t*-statistics. Panel B is similar to panel A, except that the independent variable is monthly consensus analyst recommendations prior to the return month. We first obtain monthly consensus analyst recommendation as median recommendation from the I/B/E/S summary file, where recommendation takes the values of 1 (strong buy), 2 (buy), 3(hold), 4(sell), or 5 (strong sell). "Sell recommendation" ("buy recommendation") is a binary variable that equals 1 if the consensus recommendation is higher (lower) than 3 and 0 otherwise. We standardize the dependent variables to facilitate the comparison of economic significances. ***, **, and * represent statistical significance at the 0.01, 0.05, and 0.10 level, respectively.

investors respond more positively to consensus buy recommendations than local investors.

We also estimate panel regressions of monthly cumulative abnormal return (CAR) in the month $t+1$ or $t+2$ on the analyst optimism measure in month t and analyst recommendations in month t, respectively. For brevity, the results are reported in Table IA3 of the Internet Appendix, which shows that analyst optimism and buy recommendations both negatively predict subsequent abnormal returns in our sample period.[34] These results indicate that the positive relations of nonlocal sentiment with analyst optimism and buy recommendations are result of nonlocals' naivety. This experiment therefore supports the view that local knowledge helps investors overcome their behavioral biases.

[34] Although analyst upgrades and downgrades typically produce returns in the direction of the recommendation, Jegadeesh et al. (2004) report that this is not always true of the level of the consensus recommendation, the measure used in this test.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

A more malevolent alternative is that some Stocktwits users are attempting to manipulate the price. Determining the veracity of a tweet is a difficult problem in sentiment analysis, but our results on recommended posters in panel D of Table 4 provide some suggestive evidence. It has been shown that reputation is valuable in social networks. For example, Da and Huang (2016) examine the data from Estimize.com, an online platform where users make earnings forecasts, and find that forecasts made by influential users have much stronger impact on the forecasts made by subsequent users. Our results on recommended posters show that nonlocal disadvantage is no different between the recommended posters with a reputation to protect and the full sample. This result provides suggestive evidence that the veracity of the bulk of the sample is no different from the subgroup subject to reputational constraints.

As another check for evidence of manipulation in the database, we note that Leinweber and Madhavan (2001) find a "pump and dump" strategy is by far the most common form of price manipulation, so we evaluate the possibility that the observed nonlocal disadvantage is caused by nonlocals using "pump and dump" strategy. Specifically, we examine penny stocks priced below $2 because "pump and dump" can be more profitable for penny stocks.[35] Inconsistent with the "pump and dump" explanation, we find no evidence that nonlocal disadvantage is greater in penny stocks than our sample of large stocks. Both of these manipulation tests are indirect, but they show no evidence that nonlocal tweets are positively biased in an attempt to manipulate the stock price.

5. Conclusion

This paper investigates the ability of sentiment information extracted from a social network to provide useful investing information. Because such sentiment information is increasingly provided to individual investors with little guidance on its use, investors should understand that naively following sentiment information can be dangerous for their portfolio. At the same time, this information can be valuable if it is used in a sophisticated manner. To demonstrate this, we investigate local and nonlocal posters using a unique data set of Twitter posts that cover publicly traded U.S. companies. Whereas previous studies on individual investors' local advantage focus on the abnormal returns on investors' local investments, we examine individual investors' sentiment about local and nonlocal companies.

We first examine the overall predictive ability of Twitter sentiment for stock returns and find that the sentiment in these posts exhibits significantly negative predictive ability. This finding demonstrates the danger to individual investors who blindly use social network sentiment to guide their investment

[35] We thank Jonathan Berk for this suggestion.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

decisions. When we contrast the stock return predictive ability between locals and nonlocals, we find that whereas nonlocal sentiment has a significantly negative predictive ability, local sentiment does not. The gap between locals and nonlocals, or the nonlocal disadvantage, is both economically and statistically significant. Further analyses show that nonlocal disadvantage is much larger in firms without news coverage and in firms with higher information asymmetry.

We examine the relation between nonlocal disadvantage and mispricing. Consistent with nonlocal disadvantage being driven by stock mispricing, we find that nonlocal disadvantage is much more pronounced in stocks with greater limits to arbitrage as measured by higher idiosyncratic return volatility. Additionally, nonlocal disadvantage is much larger around earnings announcement windows, consistent with mispricing being corrected by earnings news. Furthermore, long-short strategies based on the nonlocal disadvantage reveal that nonlocal disadvantage is driven by the negative returns of stocks favored by nonlocal posters, suggesting that higher nonlocal sentiment in overpriced stocks drives the nonlocal advantage.

Consistent with nonlocals chasing overpriced stocks, we find that nonlocal sentiment has a much stronger positive relation with contemporaneous stock return and trading volume than local sentiment. Our results show that nonlocal sentiment exhibits significantly stronger relations with analyst optimism and analyst buy recommendations than does local sentiment.

Our findings have implications for the rapidly growing communication available on the Internet about financial markets. Many people perceive that messages available on the Internet about the stock market simply contain noise or reflect investor sentiment unrelated to firm fundamentals. We find that, indeed, Twitter posts, on average, have a negative return predictive ability. However, we also observe an insignificant relation between local post sentiment and future returns. This finding suggests that communication available via the Internet about financial markets can contain value-relevant information, but investors must choose more sophisticated ways to analyze the data to use social network information effectively.

Appendix

A.1 Matching Tickers to PERMNOs

We use PERMNOs to identify sample firms to assist in merging the data sets. Since both the Stocktwits messages and the news articles are based on stock tickers, we create a linking file that assigns PERMNO to a TICKER date from 2008 to 2011. We first download the CRSP daily stock file from January 2008 to December 2011, and identify the first and last dates of each PERMNO-ticker pair. Then, for each calendar day from January 2008 to

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

December 2011, we assign the corresponding PERMNO to a ticker as long as the day is between the first and the last days of the PERMNO-ticker pair. We then examine the resultant matches and find that most of the PERMNO-ticker pairs are one-to-one matches for a given day, but a very small number have multiple matches between PERMNO and the ticker on a given day. We address these multiple matches as follows:

1. One PERMNO matched to two tickers: Two PERMNOs 90469 and 91501 are each matched to two tickers on some days. This is due to the change in tickers during an interim period. For example, PERMNO 90469's ticker is ARBX for most of the time during our sample period, but for the 1-month interim period from June 14, 2010 to July 12, 2010, its ticker changes to ARBXD. Therefore, our procedure of using the start and end dates assigns both the tickers ARBX and ARBXD to the PERMNO 90469 for this 1-month period. We address this issue by keeping only the valid tickers (ARBXD in the case of PERMNO 90469) for these two tickers in the subperiods.
2. One ticker matched to two PERMNOs: From 2008 to 2011, a small number of tickers each matched to two PERMNOs for either the whole period or a subperiod. We find that these cases are due to a firm issuing shares of two classes that correspond to two different PERMNOs (e.g., shares with voting power vs. shares without voting power). To address this issue, for each of these tickers, we calculate the total share volume for two PERMNOs, respectively, from 2008 to 2011, and keep the PERMNO with the larger share volume. In most cases, the share volume of one PERMNO is much larger than the other.

A.2 Collecting Twitter Users' Location Information

We identify a poster's location using the following approach:

1. For each user, we first search the user's account ID on Stocktwits.com to locate the profile page and record the user location(s).
2. If a user's Stocktwit profile page does not contain location information, we then search the account ID on Twitter.com to pull up the user profile. A small number of these users provide location information on their Twitter.com profile. Since the same account ID can correspond to different users on Stocktwits.com and Twitter.com (e.g., the account "Tony" on Twitter could be a different user than "Tony" on Stocktwits.com), we use a poster's location from Twitter.com only when we have enough evidence that the account belongs to same user as on Stocktwit.com; most often the profile picture is the same

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

in the Stocktwits profile and the Twitter profile. Only a small number of user locations are collected using this approach.

In a rare situation, a user provides more than one location. In this case, we include the user in our sample as long as one of the locations is in the continental United States. Additionally, when we identify local posts in this case, a post is considered local as long as one of the user locations is local to the company discussed in the posts (within 100 miles of the corporate headquarters). About one-third of users have available location information.

The user locations for our sample contain the state and city (county) information. We convert user locations into coordinates using http://itouchmap.com/latlong.html, which provides coordinates for the center of a city (county). We then use the coordinates to calculate distance between a user and a corporate headquarter according to Equation (1) in the main paper.

A.3 Maximum Entropy Approach

To formally describe the maximum entropy (ME) procedure, we define the following set of terms. Let $F = (f_1, \ldots f_m)$ be a set of predefined features that can appear in a post. From our previous example, the word "sell" would be a feature, and the trigram "crazy to sell" also would be a feature. Let $n_i(d)$ be the number of times that the feature f_i occurs in a post d. Thus, each post is represented by a post vector that takes the form: $\bar{d} = (n_1(d),\ n_2(d),\ n_m(d))$. Lastly, let c be a post category that takes the value of c_0 (positive, negative, or neutral). Given this set of variables, the estimate of $P(c{=}c_0|d)$ is as follows:

$$P_{ME}(c = c_0 \ \ d) = \frac{1}{Z(d)}\left(\sum_i \lambda_{i,c} F_{i,c}(d,\ \ c)\right), \tag{A1}$$

where $Z(d)$ is a normalization function, and $F_{i,c}$ is a feature category function for the feature i and for each category c defined as

$$F_i, c(d, c) = \begin{cases} 1, & \text{if}\, n_i(d) > 0 \ \text{ and } c_i = c_0, \\ 0, & \text{otherwise.} \end{cases} \tag{A2}$$

For example, this feature category function only returns a value of one if the post contains the trigram "crazy to sell" and the post is hypothesized to be of positive sentiment. $\lambda_{i,c}$ is a weighting parameter that determines the relative strength of each of the features f_i contained in a document. If the value of $\lambda_{i,c}$ is very large then the feature f_i is considered to be very strong for a specific category c_0. The weighting parameter allows us to implement Jegadeesh and Wu's (2013) finding that weighting can be an important tool in content analysis.

We implement the ME classifier by hand classifying a corpus of 2,000 twitter posts. This out-of-sample set of categorized data is referred to as

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

the training set and is used to calculate the expected values of $F_{i,c}$. Next, we use all the Twitter posts to estimate the conditional probabilities $P_{ME}(c = c_0|d)$ by calculating the maximum likelihood solution across the three different categories, while satisfying the constraint that the expected values of the feature category functions $F_{i,c}$ are equal to their training data expected values. Each post in our data set is then assigned a value of (-1, 0, 1) based on the highest conditional probability of a post being positive, negative, or neutral. We test the accuracy of this procedure by running the ME classifier on a set of 300 posts that are hand-classified. The ME classifier worked well in this out of sample test, and it was able to correctly classify 67% of all posts in the test sample. This accuracy rate is similar to the accuracy level that is achieved in other sentiment classification studies, such as Pang, Lee, and Vaithyanathan (2002).

A.4 Collection of News Stories

The news search is based on the tickers and firm names. For each stock ticker covered by Stocktwits, we collect the corresponding firm name from the CRSP monthly stock file during the sample period. We then search the news stories from Dow Jones Newswire, Reuters News, and PR Newswire from November 11, 2008 and June 10, 2011. When we search a firm, we first enter the ticker, and then pick a name from Factiva's suggested list of firm names that matches the firm's name in CRSP. We also eliminate the duplicates of news stories for a given firm. We then matched the articles to PERMNOs using the approach described in Appendix A.1. A small number of unmatched articles are outside the date ranges of CRSP for the corresponding tickers. This happens because, even when a firm is not traded in the exchange, it can still have news coverage. For example, General Motors (PERMNO 12079) stopped trading on June 1, 2009 and resumed trading on November 18, 2010, with a new PERMNO of 12369. GM's news articles during this interim period are therefore not matched to a PERMNO.

References

Agarwal, S., and R. Hauswald. 2010. Distance and private information in lending. *Review of Financial Studies* 23:2757–88.

Almazan, A., A. De Motta, S. Titman, and V. Uysal. 2010. Financial structure, acquisition opportunities, and firm locations. *Journal of Finance* 65:529–63.

Anand, A., V. Gatchev, L. Madureira, C. Pirinsky, and S. Underwood. 2011. Geographic proximity and price discovery: Evidence from the Nasdaq. *Journal of Financial Markets* 14:193–226.

Antweiler, W., and M. Z. Frank. 2004. Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59:1259–94.

Ayers, B., S. Ramalingegowda, and E. Yeung. 2011. Hometown advantage: The effects of monitoring institution location on financial reporting discretion. *Journal of Accounting and Economics* 52:41–61.

Bae, K., R. M. Stulz, and H. Tan. 2008. Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts. *Journal of Financial Economics* 88:581–606.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Baik, K., J. Kang, and J. Kim. 2010. Local institutional investors, information asymmetries, and equity returns. *Journal of Financial Economics* 97:81–106.

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

Bernile, G., A. Kumar, and J. Sulaeman. 2015. Home away from home: Economic relevance and local investors. *Review of Financial Studies* 28:2009–49.

Bradshaw, M. T., S. A. Richardson, and R. G. Sloan. 2006. The relation between corporate financing activities, analysts' forecasts and stock returns. *Journal of Accounting and Economics* 42:53–85.

Brennan, M. J., and A. Subrahmanyam. 1995. Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38:361–81.

Butler, A. W. 2008. Distance still matters: Evidence from municipal bond underwriting. *Review of Financial Studies* 21:763–84.

Chen, H., P. De, Y. Hu, and B. Hwang. 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies* 27:1367–403.

Coval, J. D., and T. J. Moskowitz. 1999. Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54:2045–73.

———. 2001. The geography of investment: informed trading and asset prices. *Journal of Political Economy* 109:811–41.

D'Avolio, G. 2002. The market for borrowing stock. *Journal of Financial Economics* 66:271–306.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic based benchmarks. *Journal of Finance* 52:1035–58.

Da, Z., and X. Huang. 2016. Harnessing the wisdom of crowds. Working Paper, University of Notre Dame.

Das, S., M. Asis, and P. Tufano. 2005. eInformation: A clinical study of investor discussion and sentiment. *Financial Management* 34:103–37.

Driscoll, J., and A. Kraay. 1998. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* 80:549–60.

Dwoskin, E. 2013. Twitter's data business proves lucrative; Twitter disclosed it earned $47.5 million from selling off information it gathers. *Wall Street Journal Online*, October 7. https://www.wsj.com/articles/twitter8217s-data-business-proves-lucrative-1381104711.

Eggers, B. D. 2014. *Social media strategies for investing*. Avon, MA: Adams Media.

Engelberg, J. 2008. Costly information processing: Evidence from earnings announcements, Working Paper, University of North Carolina.

Engelberg, J., D. R. McLean, and J. Pontiff. 2016. Anomalies and news. Working Paper, Boston College.

Feng, L., and M. S. Seasholes. 2005. Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance* 9:305–51.

Gaspar, J., and M. Massa. 2007. Local ownership as private information: Evidence on the monitoring-liquidity trade-off. *Journal of Financial Economics* 83:751–92.

Fama, E. F., L. Fisher, M. C. Jensen, and R. Roll. 1969. The adjustment of stock prices to new information. *International Economic Review* 10:1–21.

Fang, L. H., and J. Peress. 2009. Media coverage and the cross-section of stock returns. *Journal of Finance* 64:2023–52.

Hatheway, F., A. Kwan, and H. Zheng. Forthcoming. An empirical analysis of market segmentation on U.S. equities markets. *Journal of Financial & Quantitative Analysis*.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Hendershott, T, D. Livdan, and N. Schürhoff. 2015. Are institutions informed about news? *Journal of Financial Economics* 117:249–87.

Hirshleifer, D., S. Lim, and S. H. Teoh. 2011. Limited investor attention and stock market misreactions to accounting information. *Review of Asset Pricing Studies* 1:35–73.

Hong, H., T. Lim, and J. C. Stein. 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55:265–95.

Huang, A. H., A. Y., Zang, and R. Zheng. 2014. Evidence on the information content of text in analyst reports. *Accounting Review* 89:2151–80.

Irvine, P. J. 2004. Analysts' forecasts and brokerage-firm trading. *Accounting Review* 79:125–49.

Ivković, Z., and S. Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *Journal of Finance* 60:267–306.

———. 2007. Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices. *Review of Financial Studies* 20:1327–57.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48:65–91.

Jegadeesh, N., J. Kim, S. Krische and C. Lee. 2004. Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59:1083–124.

Jegadeesh, N. and D. Wu. 2013. Word power: A new approach for content analysis. *Journal of Financial Economics* 110:712–29.

Korniotis, G. M., and A. Kumar. 2011. Do older investors make better investment decisions? *Review of Economics and Statistics* 93:244–65.

Leinweber, D. J., and A. N. Madhavan. 2001. Three hundred years of stock price manipulations. *Journal of Investing* 10:1–10.

La Porta, R. J. Lakonishok, A. Shleifer, and R. Vishny. 1997. Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52:859–74.

Li, F. 2010. The information content of forward-looking statements in corporate filings - a naïve Bayesian machine learning approach. *Journal of Accounting Research* 48:1049–102.

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

Malloy, C. J. 2005. The geography of equity analysis. *Journal of Finance* 60:719–55.

Malmendier, U., and D. Shanthikumar. 2007. Are small investors naïve about incentives. *Journal of Financial Economics* 85:457–89.

———. 2014. Do security analysts speak in two tongues? *Review of Financial Studies* 27:1287–322.

Mikhail, M., B. Walther, and R. Willis. 2007. When security analysts talk, who listens? *Accounting Review* 82:1227–53.

Miller, E. 1977. Risk, uncertainty, and the divergence of opinion. *Journal of Finance* 32:1151–68.

Pang, B., L. Lee, and S. Vaithyanathan. 2002. Thumbs up? Sentiment classification using machine learning techniques. Proceedings, ACL-02 Conference on Empirical Methods in Natural Language.

Reed, A. 2013. Short selling. *Annual Review of Financial Economics* 5:245–58.

Seasholes, M. S., and N. Zhu. 2010. Individual investors and local bias. *Journal of Finance* 65:1987–2010.

Seru, A., T. Shumway, and N. Stoffman. 2010. Learning by trading. *Review of Financial Studies* 23:705–39.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019

Stambaugh, R. F., J. Yu, and Y. Yuan. 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics* 104:288–302.

———. 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70:1903–48.

Tetlock, P. C. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62:1139–68.

———. 2010. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:3520–57.

Tumarkin, R., and R. F. Whitelaw. 2001. News or noise? Internet postings and stock price. *Financial Analysts Journal* 57:41–51.

Wermers, R. 2004. Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working Paper, University of Maryland.

Downloaded from https://academic.oup.com/raps/article-abstract/8/2/293/4056137 by texas christian university user on 08 November 2019