Stock Mispricing, Hard-to-value Stocks and the Influence of Internet Stock Message Boards

Xiong, X., Meng, Y., Joseph, N. & Shen, D.

Original citation & hyperlink:
Xiong, X, Meng, Y, Joseph, N & Shen, D 2020, 'Stock Mispricing, Hard-to-value Stocks and the Influence of Internet Stock Message Boards', International Review of Financial Analysis, vol. 72, 101576.
https://dx.doi.org/10.1016/j.irfa.2020.101576

DOI 10.1016/j.irfa.2020.101576
ISSN 1057-5219

Publisher: Elsevier

International Review of Financial Analysis

Volume 72, November 2020, 101576

https://doi.org/10.1016/j.irfa.2020.101576

Stock Mispricing, Hard-to-value Stocks and the Influence of Internet Stock Message Boards

Xiong Xiong [a], Yongqiang Meng [a], Nathan Lael Joseph[b] and Dehua Shen [a]

[a]*College of Management and Economics, Tianjin University, Tianjin, PR China*

[b]*Centre for Financial and Corporate Integrity, Coventry University, Priory Street, Coventry, United Kingdom, CV1 5FB*

Acknowledgements
Early version of this paper entitled "Does internet stock message board serve as the proxy for firm-specific investor sentiment?" is present at the 2019 International Conference on Internet Finance. Helpful comments from participants are gratefully acknowledged. This work is supported by the National Natural Science Foundation of China (71532009, 71790594 and 71701150).

Corresponding author
Dehua Shen, Associate Professor of Finance
College of Management and Economics, Tianjin University
No.92 Weijin Road, Nankai District, Tianjin, 300072, China
Email: dhs@tju.edu.cn

Highlights

- The information in Internet Stock Message Boards (ISMB) postings certain characteristics that relate to investor sentiments.
- Short-term ISMB postings contain persistence in hard-to-value stocks.
- Stocks with high ISMB postings underperform in the long-term.
- ISMB postings proxy reasonably well for firm-specific investor sentiment.
- ISMB postings are associated with temporary mispricing in stocks.

Abstract

In recent years, the predictability of Internet Stock Message Board (ISMB) postings has been intensively investigated. However, the underlying mechanisms driving the ISMB postings and their influence on hard-to-value stocks remain largely unexplored. In this paper, we show that the information contained in the process underlying ISMB postings contains characteristics that are associated with investor sentiment. In particular, we show that short-term ISMB postings contain persistence in hard-to-value stocks and stocks with high ISMB postings underperform in the long-term. Our empirical findings indicate that ISMB postings proxy reasonably well for firm-specific investor sentiment and are associated with temporary mispricing in stocks.

1. Introduction

The literature posits an important connection between behavioral biases and stock valuation when information is sparse. Stocks with sparse information flows are considered harder to value. The associated behavioral biases implicitly or explicitly cause investors to make mistakes in their (subjective) valuations. The literature debates whether the observed price valuations are the outcome of behavioral biases (mispricing) or compensation for risk. Both factors may have relevance to our study.

The use of internet data to investigate investor stock trading behavior has become an important area of academic research.[1] The findings from some of these studies are however, inconsistent. For instance, using the RagingBull.com Internet Stock Message Board (ISMB), Tumarkin and Whitelaw (2001) claim that message board activity does not predict abnormal trading volume and industry-adjusted returns. However, after controlling for fundamental news of the largest 45 firms in the Dow Jones Industrial Average, Antweiler and Frank (2004) find that Internet postings associated with stock information help predict market volatility. Sabherwal et al. (2011) also claim that message board activity can help predict the price behavior of small stocks. On balance, prior studies suggest that internet message board postings associated with stock prices, contain information that predicts stock returns. As such, we focus on the information contained in ISMB postings to relate to stock mispricing and hard-to-value stocks. We motivate our study below.

Due to the unavailability of the firm-level proxies for investor sentiment, the literature relating to the impact of investor sentiment on stock price dynamics mainly relies on market-wide measurement measures.[2] Baker and Wurgler (2006) predict that investor sentiment has cross-sectional effects on stock prices when stock valuations are subjective and difficult to arbitrage. They argue that while it is difficult to distinguish between both

---

[1] Related prior studies include those by Bollen et al. (2011), Da et al. (2011), Das and Chen (2007), Dimpfl and Jank (2016), Joseph et al. (2011), Kim and Kim (2014), Leung and Ton (2015), Shen et al. (2018), Siganos et al. (2014), Sprenger et al. (2014), Wysocki (1998) and Zhang et al. (2016). These studies use several internet sources including ISMB, Google Trends, Baidu Index (News), Twitter, Facebook, and Sina Weibo.

[2] See Baker and Wurgler (2006), Da et al. (2015), Danso et al. (2019), Hribar and McInnis (2012) and Mian and Sankaraguruswamy (2012).

channels in empirical tests, stocks that are hardest to arbitrage are also difficult to value. Furthermore, the theoretical models associated with investor psychological biases (Daniel et al., 1998; 2001; Hirshleifer, 2001), predict that investor behavioral biases are more pronounced for hard-to-value stocks. Kumar's (2009) empirical evidence, for example, shows that stock and market-level uncertainty increase behavioral biases, that informed investors exploit.

In our study, we use the information contained in ISMB postings to capture the potential subjectivity in the valuations of stocks that are hard-to-value and difficult to arbitrage (Baker and Wurgler, 2007). In order to capture investor sentiment and the influence of the information contained in internet postings on price behavior, we focus on the constituent stocks in the China Securities Index (CSI) 300. These are the 300 most representative stocks in the Chinese stock market. They have the largest market capitalization values in the universe of A-listed share companies in China.[3] The index captures the overall performance of A-listed shares on both the Shanghai Stock Exchange and the Shenzhen Stock Exchange. From this sample of stocks, we generate measures that proxy for hard-to-value stocks and relate our sentiment measure to their potential mispricing.

Our study therefore investigates whether ISMB postings adequately proxy for firm-level investor sentiment of hard-to-value stocks and stocks that are difficult to arbitrage in the Chinese stock market. We use the ISMB postings as a sentiment measure, since internet postings provide additional information regarding investors' trading behavior (Sabherwal et al., 2011; Li et al., 2018b). Furthermore, in line with related studies in our area (Sabherwal et al., 2011), we move away from the use of a market-wide sentiment measure (Baker and Wurgler, 2006), to an investor sentiment measure that is specific for individual stocks.

---

3 The constituents of the CSI 300 index are stocks with good performance attributes. Such stocks are without serious financial problems or violation of laws and regulations and have few large price volatilities that may be indicative of market manipulation. In addition, strict criteria are required for inclusion in the CSI 300. See http://www.csindex.com.cn/en/indices/index-detail/000300.

We focus on the Chinese stock market for several reasons. First, both the quantity and proportion of retail investors in the Chinese stock market are larger compared to those of both developed and other emerging stock markets. Indeed, at the end of 2018, there were 146 million retail investors, along with 353 thousand institutional investors operating in the Chinese stock market.[4] Furthermore, the large number of available retail investors suggests that the sway of opinions will likely bring about a high level of consensus in stock prices, which in turn may capture investor sentiment. While the large body of retail investors suggests that there is likely to be a substantial amount of information influencing Chinese stock price movements, we do not presume that all these retail investors are active participants of our internet message board.[5] Second, Zhang et al. (2016) show that the cumulative abnormal returns of *Announcement Interpretation* associated with internet news in the Chinese stock market, reverses completely after 50 trading days. While their results support the price pressure hypothesis, they find limited evidence to support the information diffusion hypothesis.[6] An important implication of their result is that internet message boards appear to provide evidence for investor sentiment and that mispricing appears to exist in stock prices. Third, to capture retail investor sentiment, prior studies focus on retail investor behavior, including the number of opening accounts (Huang et al., 2009; Ni et al., 2015) and related internet information (Joseph et al., 2011; Kim and Kim, 2014).[7] However, the use of ISMB postings to capture the sentiment of retail investors is

[4] See the China Securities Depository and Clearing Corporation Statistical Yearbook (2018), http://www.chinaclear.cn/zdjs/editor_file/20190716171026726.pdf

[5] In a recent survey conducted by Shenzhen Stock Exchange, individual investors account for 75.1% (2018) in Chinese stock market, while the corresponding figures for U.S. equity market and London Stock Exchange are about 27% (2009) and 12.4% (2014), respectively. See: http://www.szse.cn/aboutus/trends/news/t20180315_519202.html; https://www.sec.gov/news/speech/2013-spch041913laahtm#P18_1663; https://www.ons.gov.uk/economy/investmentspensionsandtrusts/bulletins/ownershipofukquotedshares/2016.

[6] Under the price pressure hypothesis, news in the market place including internet news creates temporary buying pressure which increases abnormal returns that reverses in the very short term. The information diffusion hypothesis predicts that relevant information regarding stock fundamentals leads to abnormal returns that do not reverse in the very short term.

[7] Noise traders also exist in markets that contains retail and institutional investors and their presence can affect the relationship between sentiment and stock returns (Kumar and Lee, 2006; Baker and Wurgler, 2006). In the De Long et al. (1990)'s overlapping generation model, noise trader risk is borne

interesting in the Chinese context, since the stock market operates under a regime of price limits, T+1 trading mechanism and limited short-selling activity. Specifically, the daily 10% price limit in the Chinese stock market restricts the diffusion of new information in stock prices when the price limit is reached, and in turn, limit arbitrage activity (Chen et al., 2019). This restriction is also likely to place a cap on the strength of investor sentiment, albeit temporarily. If indeed the price pressure hypothesis holds in Chinese stocks (Zhang et al., 2016), this setting makes an interesting case for using the information contained in internet message boards as a proxy for investor sentiment. Furthermore, the uniquely T+1 trading mechanism in the Chinese stock market implies the stocks bought during a day cannot be sold the same day (Zhang, 2020). This trading system is different from other universal trading systems that allow T+0 trading, where stocks can be bought and sold the same day. Thus, the combination of price limit and the T+1 trading mechanism delays adjustment of stock prices to new positive and negative information. Zhang (2020) shows that the T+1 trading mechanism leads to negative overnight returns whereas, overnight returns in T+0 trading markets, even in the Chinese stock index futures markets, are zero or positive. The *overnight return puzzle* – tendency for overnight returns to be negative – appears to be unique to Chinese stock markets (Qiao and Dam, 2020). Fourth, under the overvaluation hypothesis, the presence of short-selling constraints leads to a divergence of opinion whereby stock prices are higher than their fundamental value (Scheinkman and Xiong, 2003).[8] As such, optimistic investors are able to buy stocks unconstrained, other than by their available funds, and pessimistic investors are unable to sell stocks they no longer have. However, unlike many developed and emerging markets, short-selling in the Chinese stock market has operated in fits and starts.[9] There is a very limited history of short-selling

---

by arbitrageurs with short-return horizon which limits their willingness to bet against noise traders. This is turn causes prices to diverge from fundamentals and for noise traders to earn profits above those of fundamentals.

[8] Using an experimental study, Hauser and Huber (2012) report that short-selling constraints systematically distort prices to diverge from fundamentals and for noise traders to earn profits above those of fundamentals such that high (low) capitalization stocks trade at lower (higher) prices relative to their fundamental value and no short-selling restrictions lead to more efficient prices.

[9] https://www.investopedia.com/ask/answers/09/short-selling-china.asp (Accessed, June 7, 2020)

activity (Li et al., 2018a) which in turn constraints diversion of investment opinions. Thus, in a stock trading environment of daily price limits, T+1 trading mechanism, restricted short-selling activity, it would be economically useful to examine the role of ISMB postings in influencing prices of hard-to-value stocks, and the associated mispricing that can occur. Indeed, the presence of T+1 trading mechanism reduces total trading volume and volatility (Guo et al., 2012), and tightens stock liquidity, such that stocks are sold at a discount at low levels of liquidity (Bian et al., 2017). However, while the presence of the T+1 trading mechanism is associated with price discounts on daily openings which contribute to overnight risk (Qiao and Dam, 2020), we suggest that the large number of retail investors in our ISMB postings will bring about a degree of consensus in stock prices. Finally, our study focuses on one of the world's most important emerging stock markets and economies. Given the large number of retail investors in the Chinese stock market, it is useful to investigate whether the US findings translate to other markets. These factors make the Chinese market an interesting environment to examination.

Our results are summarized as follows. We find evidence for short-run persistence in ISMB postings that is highest for the highest decile. Stocks in the lowest portfolio decile exhibit persistence that lasts for up to four weeks. However, the sensitivity of investor sentiment, measured by the difference between the highest and lowest decile internet postings, does not relate to stock market beta, firm size or book-to-market (BM) ratio. This suggests that our measure of sentiment does not influence proxies for risk factors including systematic risk. However, the sensitivity of short-run investor sentiment positively relates to portfolios with the highest momentum returns and highest stock turnover. As such, the effects of investor sentiment are asymmetric and occurs for measures associated with mispricing behavior. Brown and Cliff (2005) argue that limits to arbitrage may cause sentiment to be asymmetric. As such, the limited short-selling arrangements in the Chinese stock market can cause investor sentiment to be asymmetric. We also find evidence of strong price reversals.

Baker and Wurgler (2006), Berkman et al. (2012) and Hribar and McInnis (2012) claim that investor sentiment has more influence on stocks that are hard-to-value. Using several proxies of hard-to-value stocks, we find that stocks with higher volatility, larger size (capitalization), younger age, less profitability, and higher price-to-earnings ratios, have the largest number of ISMB postings. Baker and Wurgler (2007) predict that such stocks are disproportionately exposed to broad waves of investor sentiment. We suggest that the higher ISMB postings associated with larger size, for example, also occur due to the greater visibility of large stocks and higher investor attention (Wysocki, 1998). Our result suggests two important factors. First, investors do not simply ignore hard-to-value stocks. Second, they appear to be more uncertain about the value of hard-to-value stocks. We also find that for a given week, stocks in the highest portfolio decile of ISMB postings significantly underperform those in the lowest portfolio decile using holding period of up to 200 weeks. This result suggests that strong positive market sentiment drives stock prices to levels that are not sustainable, giving rise to price reversals. This result is in line with US studies that indicate investor sentiment is negatively correlated with long-run stock returns (Baker and Wurgler, 2006; Da et al., 2015).

Our results contribute to prior results in the following ways. First, our results indicate that ISMB postings contain information that reasonably proxies for investor sentiment. However, the observed level of sentiment is asymmetric and specifically relates to return momentum and stock turnover at the highest portfolio quartile. Momentum and stock turnover are often associated with behavioral biases that lead to mispricing (Baker and Wurgler, 2006). We contribute to the prior studies of Danso et al. (2019), Hribar and McInnis (2012) and Mian and Sankaraguruswamy (2012), by showing that ISMB postings capture specific elements of retail investors' market behavior. Second, we add to the evidence by showing that information in ISMB postings predicts stock returns. The predictive effects are more pronounced at higher levels of investor sentiment, with the possibility that they are associated with price overreaction. Our results show that proxies of hard-to-value stocks become useful predictors of stock price behavior (Antweiler and Frank, 2004). Our paper also contributes to the literature associated with investor inattention and the release of public information. Theoretical models of investor behavior and information asymmetry suggest that limited attention of investors is associated with price underreaction (Hirshleifer, 2001; Hirshleifer et al., 2009; Hirshleifer et al., 2013). Indeed, investors react to a narrower set of stocks that just caught their attention (Odean, 1999). We show that lower decile-sized stocks have fewer ISMB postings. Yet, the lower deciles of ISMB postings contain positive and significant returns, thus providing evidence of underreaction. Our paper differs from the papers by Sabherwal et al. (2011) and Ackert et al. (2016) that indicate influential investors are unable to manipulate the stock market with a pump and dump trading strategy.

We structure the remaining sections of the paper as follows. Section 2 reviews the related literature and develops our hypotheses. Section 3 describes the data used in the study. Section 4 presents the empirical findings and section 5 summarizes and concludes the paper.

2. Literature review and hypotheses development

This section provides a review of prior studies. We focus on the literature associated with investor behavioral biases and their relationship with investor sentiment. We briefly review the sources of investor behavioral biases including the literature on internet message boards.

2.1 Media information and retail investor sentiment

The behavioral and cognitive biases of investors convey information about mispricing. Barber and Odean (2008) show that individual investors are more likely to be net buyers of attention-grabbing stocks compared to institutional investors. Baber et al. (2009) report that individual investors herd, and that small trade order imbalances reliably predict returns in the opposite direction. Indeed, Yao et al. (2014) document differences in herding behavior of Chinese A and B stocks. Using consumer confidence surveys, Qiu and Welch (2004) find that investor sentiment plays an important role in financial markets. Moreover, Gao and Yang (2018) find that high investor sentiment and high investor trading strengthen the positive relation between sentiment-returns and behavior-returns in Chinese future market. Indeed, Mian and Sankaraguruswamy (2012) argue that sentiment-driven mispricing of earnings contributes to the general mispricing of stocks. Kumar and Lee

(2006) report that systematic retail trading explains return movements for stocks with high retail concentration, especially when stocks are costly to arbitrage.

Traditional media, e.g., business press, and social media, e.g., internet messaging boards, are often used to capture stock price behavior. Traditional measures of investor sentiment adopt a market-wide approach (Baker and Wurgler, 2006; Huang et al., 2009). Da et al. (2015) use daily internet search volume to construct an index of market-level sentiment. Joseph et al. (2011) proxy investor sentiment using online ticker searches. They report that online searches predict abnormal returns and trading volumes and that online search sensitivity is positively related to the difficulty of arbitraging stocks. Da et al. (2011) measure investor attention using the search frequency in Google, i.e., Search Volume Index (SVI). They report that increases in SVI predict higher stock prices over the following two weeks, which reverse within a year. Furthermore, Dimpfl and Jank (2016) report that search queries Granger-cause volatility, and increases in internet searches are associated with greater volatility the following day. Solomon (2012) finds that more media coverage including business press coverage increases announcement returns. Using 2.2 million articles from 45 US newspapers, Hillert et al. (2014) find that stocks with high firm-specific media coverage exhibit stronger momentum. Zhang et al. (2016) also report that abnormal returns and excess volume trading are positively related with internet news on the event date.

In recent years, a body of literature has developed around the effects of social media sentiment and internet postings. Prior studies show that social media sentiment and internet postings are strongly associated with abnormal returns and trading volume (Tumarkin and Whitelaw, 2001; Antweiler and Frank, 2004; Bollen et al., 2011; Sprenger et al., 2014; Siganos et al., 2014; Leung and Ton, 2015). Wysocki (1998) uses the data from Yahoo! Message boards and finds that cumulative volume is highest for stocks with extreme past returns and accounting performance, and stocks with high market capitalization, high price-earnings and volatility, and high market-to-book ratios. Li et al. (2018b) report that Chinese ISMB convey firm-specific information in relation to idiosyncratic volatility.

Overall, prior studies suggest the investor sentiment influences stock mispricing (Mian and Sankaraguruswamy, 2012; Huang et al., 2016). While prior studies use different approaches to measure investor sentiment, internet and media data have become important sources for capturing stock mispricing and investor sentiment (Joseph et al., 2011; Da et al., 2011; Hillbert et al., 2014). Using Chinese ISMB, we therefore predict:

H_1: Investor sentiment based on internet message boards positively relates to mispricing and stock price reversals.

2.2 Hard-to-value firms

The literature documents that investor sentiment drives short-run mispricing and long-run return reversal. Difficult decisions, especially those with delayed feedback or noise are often associated with behavioral biases (Kahneman and Tversky, 1973). The theoretical models of Daniel et al. (1998; 2001) and Hirshleifer (2001) formalize this psychological concept in terms of financial investment decisions. Their models predict that investor behavioral biases are more pronounced for hard-to-value stocks. Daniel et al. (1998, 2001) suggest that return predictability is stronger in stocks with more uncertainty, since investors tend to be overconfident about the value of hard-to-value stocks. Hong and Stein (1999) specify no particular psychological bias for their predictions. However, their slow information diffusion model predicts that small capitalization stocks experience slower information diffusion than large capitalization stocks, due to higher market friction and limited arbitrage opportunities. Lakonishok et al. (1994) argue that value stocks yield higher returns due to suboptimal investor behavior rather than the riskiness of the investment strategy. Baker and Wurgler (2007) predict that low capitalization stocks, younger stocks, unprofitable stocks, high-volatility stocks and non–dividend paying stocks are likely to be more sensitive to investor sentiments. Investor sentiment is more pronounced for hard-to-value stocks (Baker and Wurgler, 2007). Mian and Sankaraguruswamys (2012) and Hribar and McInnis's (2012) results are in line with these predictions. Furthermore, Mian and Sankaraguruswamys (2012) report that stock prices are more sensitive to good (bad) earnings news when sentiment is higher (lower), than when sentiment is low (high). Low liquidity is also associated with hard-to-value stocks (Datar et al., 1998). Given the above evidence, we predict:

H_2: Investor sentiment positively relates to hard-to-value stocks.

In line with prior studies, we measure hard-to-value stock using return volatility, firm size, firm age, profitability, and price-to-earnings ratio.

3. Data and variables

The sample contains all the constituent stocks of the CSI 300 Index. The CSI 300 Index, was introduced on April 8th, 2005. It is the first stock market index introduced by both the Shanghai Stock Exchange and Shenzhen Stock Exchange. We use the information associated with stock postings on the Guba Eastmoney message board to measure investor sentiment. This is because the Guba Eastmoney message board is the largest and most influential message board in China.[10] The Guba Eastmoney leads the industry by containing more data on effective browsing time, core network traffic and average daily number of visits. Indeed, Research Consulting rates Guba Eastmoney as the website with the most effective browsing time and average daily number of visits.[11] A message board is set up for each stock on the Guba Eastmoney message board. Millions of investors express their views and collect information on investment tips and corporate financial information, on this message board. Guba Eastmoney is commonly used in prior studies associated with Chinese ISMB postings (Huang et al., 2016; Ackert et al., 2016; Li et al., 2018b).

Our sample period begins in January 2014 and ends in December 2018. All stock price data and financial data are taken from the China Stock Market & Accounting Research (CSMAR) database. Stock prices are adjusted for stock splits and, stock and cash dividends. The non-overlapping weekly ISMB postings for stock i during week w is defined as the sum of the daily ISMB postings, beginning, 00:00 on Wednesday of week w and ending at

[10] Of the three popular sites, Guba Eastmoney (http://guba.eastmoney.com), Guba Hexun (http://guba.hexun.com) and Istock JRJ (http://istock.jrj.com), Guba Eastmoney is consistently ranked first based on investor visits and page views.

[11] http://report.iresearch.cn/content/2016/11/264957.shtml. Accessed June 9, 2020.

24:00 on Tuesday of the next week, i.e., week *w*+1. The non-overlapping stock return for week *w* is the compounded daily (close-to-close) return over the period beginning on Wednesday of week *w* and ending on Tuesday of week *w*+1. This approach is in line with prior studies (Lehmann, 1990; Baber et al., 2009; Aboody et al., 2018).

To examine the regularity of the ISMB postings, we first rank all stocks in the sample for each week *w* in ascending order, according to the weekly postings, and then partitioned the stocks into deciles. We use proxies that are associated with investor sentiment (Baker and Wurgler, 2006), stock information diffusion (Hong and Stein, 1999; Hong et al., 2000), as well as proxies associated with hard-to-value stocks (Kumar, 2009). We show the definitions of our variables in Appendix I.

4. Empirical Results

This section presents our empirical results. First, we present the descriptive statistics. We then focus on short and long-term persistence and the results of hard-to-value stocks.

4.1 Descriptive statistics

Panel A of Table 1 shows the average weekly ISMB postings. The lowest decile has an average weekly posting of 8.03 compared to 445.57 for the highest decile. Average weekly total returns monotonically increase from –0.14% in the lowest decile to 2.55% in the highest decile. These results tentatively suggest that increases in average weekly postings are positively correlated with increases in weekly total returns. Thus, average weekly postings may have information that influences weekly stock returns.

Apart from firm age and BM, the average weekly values of our measures steadily increase as average weekly postings increase (Panel A). The trend in firm age represents upward increases in line with weekly decile postings. However, at the highest deciles, especially deciles 9 and 10, the values of firm age decrease. BM also has a downward trend, following decile 4. Even so, the variations in the deciles in terms of firm age and BM are not substantial in magnitude. Indeed, if BM captures variation in growth and distressed stocks, and firm age captures survival risks, the variations across the deciles on these measures are very narrow.

Panel B of Table 1 shows both the Spearman rank correlation, r_s and Pearson correlation, r_p coefficients. As expected, the correlation coefficients for ISMB postings with firm age and BM are negative and significant (*p*-value ≤ 0.05). ISMB postings are positively correlated with total return, beta, return volatility, firm size, profitability, price-to-earnings ratio, momentum and stock turnover (*p*-value ≤ 0.01). The correlation coefficients in terms of ISMB postings are largest using stock turnover (r_s= 0.3956, *p*-value ≤ 0.01; $r_p =$ 0.4208, *p*-value ≤ 0.01). Correspondingly, the correlation coefficients in terms of ISMB postings are lowest (in absolute value when significant), using BM ratio (r_s= –0.0071, *p*-value ≤ 0.05; r_p= –0.0215, *p*-value ≤ 0.01). Stock turnover and return volatility have the largest correlations (r_s= 0.6245, *p*-value ≤ 0.01; r_p= 0.5668, *p*-value ≤ 0.01). In the case of closed-end funds, Brown (1999) shows that unusual periods of individual investor sentiment are associated with intense periods of market volatility that occur when the market is opened. Given also the positive correlations with ISMB postings, the evidence suggests important preliminary relationships among our measures. We note however, that the significant correlation coefficients do not imply causality.

<Insert Table 1 Here>

4.2 Short-run persistence and ISMB positing and total stock returns

To examine the short-run persistence in the ISMB postings, we calculate the subsequent average postings from week *w*+1 to week *w*+4. Panel A of Table 2 shows that for the lowest decile, average weekly postings monotonically increase from 25.64 in week *w*+1 to 34.14 in week *w*+4. In contrast, for the highest decile, average weekly postings monotonically decrease from 348.02 in week *w*+1 to 249.26 in week *w*+4. Therefore, the lowest average weekly postings are in the lowest deciles and the highest average postings are in the highest decile. However, unlike the lowest decile, average weekly postings in the highest decile decrease as the week length increases, whereas for the lowest decile, average weekly postings increase. Thus, the difference between the average weekly postings of the highest decile and lowest decile – ISMB sensitivity – decreases as the week length increases.

Even so, the ISMB sensitivity is statistically significant, using both the standard *t*-test and the non-parametric Wilcoxon signed-ranks test (p-value $\leq$ 0.01). Notice that even if the trends in deciles 1 and 10 postings are in the opposite direction, the increases in decile 1 postings are at a much slower pace compared to the decreases in decile 10 postings. This result suggests a weakening in the degree of persistence in the long-run. As such, if ISMB postings contain information about investor sentiment, we should expect to see short-term reversal in returns. However, the overall results suggest that ISMB postings may capture short-run persistence in stock returns.

<Insert Table 2 Here>

We now relate the average weekly ISMB postings to the average weekly total returns. First, we rank all stock returns in each week w, according to the week's ISMB postings and then partitioned the associated stock returns. Panel B of Table 2 shows strong variation in the magnitude of average weekly total stock returns across the deciles of average weekly ISMB postings. For the lower deciles, average weekly total returns are positive and significant for week w+1 to week w+4 of decile 1, week w+1 to week w+3 of decile 2, and week w+1 to week w+2 of decile 3 (p-value $\leq$ 0.10). Notice that as the deciles increase, fewer weeks contain significant total returns and they do so at reducing week length. As such, the lowest decile, i.e., decile 1, contains more weeks with positive and significance returns. This evidence diminishes as for longer horizons and as the deciles increase. Panel B also shows that there are pockets of significant total returns at week w+1 of deciles 5 and 6, and week w of deciles 8 to 10 (p-value $\leq$ 0.10). No other weeks or deciles contain significant total returns. Thus, if ISMB postings represent investor sentiment, its effects are in the lowest decile, giving rise to asymmetric effect. While the differences in the positive total returns in decile 1, over week w+1 to week w+4 are not substantial, the evidence suggests underreaction to the information contained in the ISMB postings.

Theory predicts that extreme growth and distress stocks are likely to have subjective valuations. Such stocks are also more difficult to arbitrage and are likely to be more strongly affected by investor sentiment (Baker and Wurgler, 2007). Our results indicate that low levels of investor sentiment are associated with positive returns. We find no evidence that high levels of sentiment are positively correlated with higher returns in the short term. In fact, at higher levels of sentiment, the returns are insignificant, except for week *w*. If our results are driven by sentiment, then they are consistent with the theory at the lower deciles, particularly, decile 1, meaning that such stocks have subjective valuations compared to those in the middle and higher deciles. Thus, over optimistic (pessimistic) investors undertake investment decisions that drive stock prices above (below) their fundamental values, leading to price reversals. This feature is consistent with many underreaction behavioral models of psychological biases. Indeed, Brown and Cliff (2005) report that high levels of sentiment are associated with lower returns over the next 2 to 3 years. Thus, our analysis suggests that the relation between average weekly ISMB postings (our sentiment measure) and average weekly stock returns mainly relates to periods of low sentiment (low ISMB postings) and that stock returns in higher deciles are unaffected by investor sentiment. Our results are partly in line with those of Baker and Wurgler (2006). They report that low investor sentiment is associate with higher returns for stocks in the extreme tails of the distribution, compared to those in the middle of the distribution.

4.3 Short-run persistence and stock characteristics

It is well-known that variation is stock returns are associated with differences in stock characteristics. Indeed, prior studies show that small stocks generate higher returns than large stocks (Banz, 1981; Fama and French, 1992). If indeed, investors are reluctant to hold small stocks because of insufficient information and/or behavioral biases, then increases in internet postings would likely reduce some of these biases leading to quicker price reversals. For example, increases in ISMB postings would likely increase the rate of information diffusion in small stocks making them easier to arbitrage (Hong and Stein, 1999). Moreover, Aboody et al. (2018) show that short-term return persistence is stronger in hard-to-value stocks and argue that the observed persistence may be associated with investor sentiment.

To investigate this issue, we relate the average weekly ISMB postings to a set of stock characteristics. The stock characteristics include stock beta, firm size, and BM ratio. These measures are associated with models that relate to priced factors (Fama and French, 1992; Datar et al., 1998; Avramov and Chordia, 2006). Systematic risk measures, such as beta, are also used to distinguish between the effects of rational and mispricing behavior (Fama and

French, 1992; Lakonishok et al., 1994).[12] We also include momentum returns and stock turnover in the analysis since prior studies suggest that they may explain investor sentiment and degree of information diffusion (Hong and Stein, 1999; Hong et al., 2000). The theory does not suggest a direct connection between momentum and hard to value stocks or stocks that are difficult to arbitrage (Baker and Wurgler, 2006). Indeed, stock return momentum is considered a control for the effects of underreaction/overreaction (Hong and Stein, 1999). That is, underreaction follows from slow information diffusion whereas, overreaction occurs when stocks are overbought or oversold due to behavioral biases, rather than fundamentals. Stock turnover also proxies for liquidity. Indeed, under short-selling constraints, high liquidity depicts the condition that markets are dominated by irrational investors such that stocks are overvalued (Baker and Stein, 2004). Amihud and Mendelson (1986), Chordia et al. (2000; 2001) show that market-wide liquidity measured by narrow bid-ask spread and high stock turnover predict lower stock returns in the cross-section of individual firms.

Table 3 shows the results based on our measures. The difference between the high and low deciles can be regarded as the sensitivity of sentiment to changes in the associated stock characteristic across the deciles. As such, if variation in the average weekly postings are associated with variation in the stock characteristics, it can be argued that variation in sentiment is associated variation in firm characteristics in terms of the highest and lowest deciles.

<Insert Table 3 Here>

Across the panels in Table 3, average weekly postings steadily increase from w+1 to w+4 in decile 1, whereas average weekly postings steadily decrease from w+1 to w+4 in decile 10. The result holds except for BM ratio (Panel C). This result is generally in line with the results in Panel A of Table 2. The differences across the deciles in Table 3 are highly significant (p-value $\leq$ 0.01), using the standard t-test. While the values of the stock characteristics at decile 10 tend to be larger than those of decile 1, the differences in the deciles based on beta, firm size and BM ratios are insignificant (p-value $\geq$ 0.10). The case

---

[12] While our study emphasizes mispricing in line with prior studies (Baker and Wurgler, 2006; Aboody et al., 2018), we acknowledge that there is substantial debate in the literature as to whether such anomalies are due to mispricing or whether they represent compensation for risk. For example, Lakonishok et al. (1994) attribute the contradiction between value and growth stocks to mispricing whereas, Fama and French (1992; 1995) and Penman and Reggiani (2013) attribute this difference to compensation for risk.

of beta requires some further interpretation. Variation in beta values tends to be associated with time variation in rational market-wide risk premia or time-varying cross-sectional variation is systematic risk. In the case of risk premia, investors would require compensation for risk. Variation in beta tends to associated with rational behavior, rather than behavioral biases and mis-pricing. Compensation for risk would require time-variation in beta, according to stock characteristics such as size, firm age, among others (Baker and Wurgler, 2006). Given the results for firm size and BM ratios in Panels B and C respectively, and the tendency for differences in the deciles of beta, firm size and BM ratios to be insignificant, we suggest investor sentiment and the associated mispricing that would follow do not explain our results. Our evidence therefore casts doubt on the irrational pricing hypothesis and the ability of investor sentiment to influence these measures. Prior studies indicate that small stocks with low analysts following have slower information diffusion (Hong et al., 2000). While we have not controlled for the effects of analysts following, the level of internet postings is unassociated with changes in beta, size and BM values. Of course, it is possible that our return horizon is too short to capture variation in our stock characteristics. We address this issue using panel regressions in a subsequent section.

Jegadeesh and Titman (1993) and Conrad and Kaul (1998) report that momentum strategies generate profitable returns over a 3 to 12 month horizon. If ISMB weekly postings capture investor sentiment, then they should relate to the level of momentum. Since movements in market-wide liquidity, including stock turnover, predict aggregate returns (Amihud, 2002; Jones, 2002), we also expect the level of internet postings to relate to stock turnover.

Panel D of Table 3 shows that except for quartile 4, the average weekly postings are relatively stable for both decile 1 and decile 10 in terms of our momentum measure. The differences between the momentum value of the highest and lowest decile are only significant for quartile 4 (p-value ≤ 0.01). We find a related result for stock turnover in quartile 4 (Panel E). Thus for both stock characteristics, the effect is asymmetric Both results support short-run sentiment according to $\mathbf{H_1}$, but in terms of return momentum and stock turnover.

Taken together, the level of internet postings positively relates to higher momentum and stock turnover, but the effect is asymmetric. The asymmetric response is supported in the behavioural theories (Daniel et al., 1998). Hong et al. (2000) provide related evidence for momentum and price reversals.[13] Thus, excessive optimism is likely to be associated with high stock turnover that may prevent rational investors from limiting the stock prices to levels that are consistent with fundamentals. The T+1 trading mechanism and limited short-selling in the Chinese stock market restrict speedy stock price corrections following stock price increases. While the effects of the T+1 trading mechanism and price limits may not be associated with illiquidity, the constraints imposed by these restrictions are likely to cause stock prices to underreact to new information. US evidence indicates that illiquidity makes it more difficult for stocks to trade and leads to low adjustment in prices following liquidity shocks (Bali et al., 2014).

4.4 Short-run persistence and hard-to-value stocks

Aboody et al. (2018) argue that sentiment plays an important role in the valuation of hard-to-value stocks. If ISMB postings capture investor sentiment, then we should find stronger support for investor sentiment in hard-to-value stocks. Hard-to-value stocks include stocks that are young, small in size, stocks with high volatility, extreme growth or depressed values (Baker and Wurgler, 2006).

To test the effects of investor sentiment on hard-to-value stocks, we use return volatility, firm size, firm age, profitability and price-to-earnings ratio as hard-to-value measures. We rank all the stocks in ascending order in week *w*, based on the hard-to-value proxies and partition them into quartiles. The stocks in each quartile are further divided into deciles in ascending order, according to ISMB postings. We emphasis the results for week *w*+1 ISMB postings, to save space. The untabulated results are available from the

---

[13] Specifically, Hong et al. (2000) find that the effects of analysts coverage on past loser stocks are more pronounced than their effects on past winner stocks such that loser stocks with poor analyst coverage react more slowly to bad news compared to good news. Campbell and Hentschel (1992) develop an asymmetric model that partly explains the negative skewness and excess kurtosis in returns. They show that volatility feedback has strong effects during periods of high volatility.

authors.

Table 4 shows that the weekly average postings for both decile 1 and 10 of week w + 1, increase as the quartile increases. The differences between the highest and lowest deciles postings are highly significant (p-value ≤ 0.01). Differences between quartile 4 and quartile 1 of the proxies are also significant. There is a higher concentration of the postings in quartile 4, particularly for decile 10, as well as a tendency for the t-ratios to be larger. Based on the differences between quartile 4 and quartile 1, stocks with higher volatility, larger size and higher price-to-earnings ratio have higher postings. This does not necessarily suggest they are more prone to behavioral biases, since for example, larger stocks are more likely to gain investors' attention, in line with Wysocki's (1998) findings. The differences between quartile 4 and quartile 1 for firm age and profitability are negative and significant. As such, firms with younger age and lower profitability have fewer postings suggesting they are more prone to behavioral biases. These stocks have slower information diffusion. The evidence is consistent with $\mathbf{H}_2$, indicating that firms with younger age and lower profitability are harder to value.

<Insert Table 4 Here>

4.5 Long-run portfolio returns

This section presents our tests for long-run price reversals. We use two estimation approaches. First, we construct portfolios to determine the abnormal returns from a buy and hold strategy. If large firms are less affected by investor behavioral biases than small firms, then using value-weighted portfolio will distort the results. Therefore, our portfolios are equally weighted. Next, we employ instrumental variables-general method of moments (IV-GMM) to estimate long-run performance. The regression approach allows us to capture conditional effects in our measure.

4.5.1 Long-run returns reversals

Barber et al. (2009) report that individual investors herd and that stocks that are heavily bought one week earn higher returns in the subsequent week whereas, stocks that are heavily sold, earn poorer returns. Furthermore, Hvidkjaer (2008) shows that stocks favored by investors in one month underperform those that were out of favor one month after portfolio formation and for up two years after. Baker and Wurgler (2006) argue that subjectivity in stock valuations makes certain stocks more prone to speculation. Such speculative activities are influenced by investor sentiments (Baker and Wurgler, 2006). If ISMB postings capture investor sentiment, then stocks with high short-run postings are likely to underperform those with low short-run postings in the long-run.

To examine long-run stock performance, we follow Aboody et al.'s (2018) approach. First, we form three equal-weighted portfolios after ranking all the stocks according to their ISMB weekly postings in week *w*. The stock returns are ranked each week *w* according to their ISMB weekly postings. We then partitioned the returns into deciles. The first long portfolio is associated with the stocks in the lowest decile. The second long portfolio is associated with the stocks in the highest decile. Finally, we construct a long–short portfolio consisting of highest decile and short the portfolio in the lowest decile to generate a third portfolio. The use of the third portfolio allows us to capture characteristics that may be associated with investor sentiment. For each of the three portfolios, we calculate the cumulative buy-and-hold raw returns (BHARs) for week *w* to week 5, for up to week *w* to week 200 (approx. 4 years).

Table 5 shows the long-run performance of the three portfolios. The table shows that both the first and second portfolios earn positive BHARs for all investment horizons up to week *w* to 200. The returns are highly significant (p-value ≤ 0.01). Specifically, the first portfolio of decile 1 has a return of 1.62% in week *w* to 5 which increases to 19.54% in week *w* to 200. The difference between the returns of week *w* to 5 of decile 1 and *w* to 5 of decile 10 is positive and highly significant (p-value ≤ 0.01). Except for week *w* to 5, the returns of the second portfolio i.e. decile 10 are lower than those of the first portfolio, i.e., decile 1. As such, starting from week *w* to 10 and until week *w* to 200, the first portfolio (decile 1) outperforms the second portfolio at all investment horizons. The long–short portfolio therefore generates negative and significant returns for all horizons after and including week *w* to 20 (p-value ≤ 0.05). At very long return horizons, i.e., after the *w* to 150 horizon, there is evidence of a small decline in the performance of both the first and second portfolios. The results provide strong evidence for return reversals shortly (five weeks) after the first period of portfolio formation. The evidence in line with Hvidkjaer's (2008) results. Figure 1 shows the plots for all our investment horizon. The plots are consistent with short-run reversals in returns and a tendency for the long–short portfolio to generate negative and significant BHARs. This evidence supports **H_2**. Figure 1 also suggests that in the very long-term, the performance of the portfolios are likely to converge.

<Insert Table 5 Here>

<Insert Figure 1 Here>

4.5.2 IV-GMM estimates

We further investigate the influence of firm characteristics on the performance of the long–short portfolio. To do this, we use IV-GMM regressions as this estimation method allows us to deal with potential endogeneity concerns (Huang et al., 2018). Endogeneity may be a concern in our estimations since portfolio returns and firm characteristics may be endogenous. For example, if increases in investor sentiment are associated with increases in trading volume, causing certain stocks to become more popular (Antweiler and Frank, 2004; Sabherwal et al., 2011), then using our buy-and-hold portfolio returns as the dependent variable may give rise to estimation problems. We specify our IV-GMM regression as follows:

$$\begin{aligned} \mathrm{r}_{w+i} = \alpha_0 + \beta_1 Beta_w + \beta_2 Return\_volatility_w + \beta_3 Firm\_size_w + \\ \beta_4 Firm\_age_w + \beta_5 Profitability_w + \beta_6 Price - to - earnings_w + \\ \beta_7 BM_w + \beta_8 Momentum_w + (Stock\_turnover_w = \\ Instrumental\ \ variables) + \varepsilon_w. \end{aligned} \tag{1}$$

In Eq. (1), α_0 is the intercept term. r_{w+i} denotes the buy-and-hold abnormal return (BHAR) based on the long minus short portfolio from week w+1 to week w+i, i.e., (i =5, 10, 15, 20, 25 … 200, as in Table 5). If we find that the high decile portfolios underperform the low decile portfolios, this finding provides evidence for mispricing which we relate to the effects of investor sentiment. This argument is in line with Baker and Wurgler's (2006) result that stocks that are attractive to speculators underperform in the following 12 months, when market-wide sentiment is high. The explanatory variables in Eq. (1) have been defined before. Stock turnover is instrumented in the IV-GMM regression. The $\beta's$ are the regression coefficients and ε_w is the error term. We use the producer price index (PPI), corporate goods price index (CGPI) and Li Keqiang index as instrumental variables.

Table 6 shows the results based on Eq. (1). The Table 6 shows very few specification violations. Given the number of IV-GMM regressions we run, the number of violations appear reasonable. Of course, finding suitable instrumental variables is often a trial and error undertaking. The table shows that the *C* and *J* statistics are largely insignificant. Specifically, the *C* statistic indicates a general absence of endogeneity. The *J* statistic indicates that we should have little concern for over-identification. Using the *F*-statistic, only three of our regressions have weak instruments (p-value $\geq$ 0.10). Notice that the violations occur in the medium (week w+40) to long-term (week w+200). Thus, overall, we have few concerns over model misspecification, given the number of regression estimations.

Except for beta, Table 6 shows strong evidence of reversals. Most of the stock characteristics are significant in the short-term. Stock turnover and Firm age have positive coefficients for up to week w to 40 (p-value $\leq$ 0.10). These variables have insignificant coefficients after that period. Return volatility, Firm size and Profitability have negative and significant coefficients for up to week w to 40; up to week w to 20 in the case of Profitability (p-value $\leq$ 0.10). The negative coefficients indicate that increases in these stock characteristics are associated with decreases in BHARs. Thus, stocks in the highest deciles based on these characteristics, underperform those in the lowest decile in terms of these measures. As such, these stocks are overpriced leading to overreaction. Investor sentiment can contribute to this effect. Beta has positive coefficients at almost all intervals. Beta tends to be insensitive to the influence of investor sentiment. Momentum has negative and significant coefficients for up to week w to 80 (p-value $\leq$ 0.10). The negative momentum arises because the second-decile 10 portfolio underperforms relative to our first-decile 1 portfolio. The negative momentum coefficients last for up to 20 months. The negative momentum may be due to the tendency for bad news to diffuse slowly, when investors are optimistic about market returns (Antoniou et al., 2013). Jegadeesh and Titman (1993) and Conrad and Kaul (1998) report that momentum strategies generate profitable returns over a 3 to 12 months horizon. Their results are based on the US stock market where a T+0 trading mechanism operates; there are no short-selling constraints in this market. This difference in the trading arrangements may explain the longer period of momentum returns for Chinese stocks. Overall, most of our firm characteristics indicate the tendency for price reversals.

<Insert Table 6 Here>

5. Conclusion

We investigate the suitability of the number of ISMB weekly postings to proxy for investor sentiment. We relate the weekly average postings to stock returns and other firm characteristics. We find that ISMB postings are a reasonable proxy for capturing investor sentiment. However, we find that short-run sentiment only appears to be present in high turnover and high momentum stocks. We also find evidence that positive and significant BHARs quickly reverse by about five weeks after portfolio formation. The reversal lasts for up to four years. Socks with the highest ISMB postings underperform in the long-run. Our evidence on long-run portfolio performance confirm our results.

Our results add to prior work on information diffusion in stock prices (Hong et al., 2000). We also add to the literature on the effects of investor sentiment on hard-to-value stocks (Baker and Wurgler, 2006). While some of our results are in line with those of US studies, we find that the sensitivity of short-run investor sentiment positively relates to portfolios with the highest momentum returns and the highest stock turnover. In this case, the effect is asymmetric where limits on arbitrage conditions in the Chinese market can give rise to this effect.

References

Aboody, D., Even-Tov, O., Lehavy, R., Trueman, B. (2018). Overnight returns and firm-specific investor sentiment. *Journal of Financial and Quantitative Analysis*, 53, 485-505.

Ackert, L. F., Jiang, L., Lee, H. S., Liu, J. (2016). Influential investors in online stock forums. *International Review of Financial Analysis*, 45, 39-46.

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

Amihud, Y., Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17, 223-249.

Antoniou, C., Doukas, J. A., Subrahmanyam. A. (2013). Cognitive dissonance, sentiment, and momentum. *Journal of Financial and Quantitative Analysis*, 48, 245-275.

Antweiler, W., Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, 59, 1259-1294.

Avramov, D., Chordia, T. (2006). Asset pricing models and financial market anomalies. *Review of Financial Studies*, 19, 1001–1040.

Baker, M., Stein, J. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7, 271-299.

Baker, M., Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61, 1645-1680.

Baker, M., Wurgler, J. (2007). Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives*, 21(2), 129-151.

Bali, T.G., Peng, L., Shen, Y., Tang, Y. (2014). Liquidity shocks and stock market reactions. *Review of Financial Studies*, 27, 1434–1485.

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics* 9, 3–18.

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

Barber, B. M., Odean, T., Zhu, N. (2009). Do retail trades move markets? *Review of Financial Studies*, 22, 151-186.

Berkman, H., Koch, P. D., Tuttle, L., Zhang, Y. J. (2012). Paying attention: overnight returns and the hidden cost of buying at the open. *Journal of Financial and Quantitative Analysis*, 47, 715-741.

Bian, J., Su, T., Wang, J. (2017). Non-marketability and one-day selling lockup. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2563312 (Accessed June 7, 2020).

Bollen, J., Mao, H., Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of*

Computational Science, 2, 1-8.

Brown, G. W. (1999). Volatility, sentiment, and noise traders. *Financial Analysts Journal*, 55, 82-90.

Brown, G. W., Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business*, 78, 405-440.

Campbell, J.Y., Hentschel, L. (1992). No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31, 281-318.

Chen, T., Gao, Z., He, J., Jiang, W., & Xiong, W. (2019). Daily price limits and destructive market behavior. *Journal of Econometrics*, 208(1), 249-264.

Chordia, T., Roll, R., Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics*, 56, 3-28.

Chordia, T., Roll, R., Subrahmanyam, A. (2001). Market liquidity and trading activity. *Journal of Finance*, 56, 501-30.

Conrad, J., Kaul, G. (1998). An anatomy of trading strategies. *Review of Financial Studies*, 11, 489-519.

Da, Z., Engelberg, J., Gao, P. (2011). In search of attention. *Journal of Finance*, 66, 1461-1499.

Da, Z., Engelberg, J., Gao, P. (2015). The sum of All FEARS investor sentiment and asset prices. *Review of Financial Studies*, 28, 1-32.

Daniel, K. D., Hirshleifer, D., Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance*, 56, 921-965.

Daniel, K., Hirshleifer, D., Subrahmanyam, A. (1998). Investor psychology and security market under- and over-reactions. *Journal of Finance*, 53, 1839–1886.

Danso, A., Lartey, T., Amankwah-Amoah, J., Adomako, S., Lu, Q., Uddin, M. (2019). Market sentiment and firm investment decision-making. *International Review of Financial Analysis*, 101369.

Das, S. R., Chen, M. Y. (2007). Yahoo! for Amazon: sentiment extraction from small talk on the web. *Management Science*, 53, 1375-1388.

Datar, V.T., Naik, N.Y., Radcliffe, R. (1998). Liquidity and stock returns: an alternative test. *Journal of Financial Markets*, 1, 203–219.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98, 703-738.

Dimpfl, T., Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, 22, 171-192.

Fama, E.F., French, K.R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47, 427–465.

Fama, E.F., French, K.R. (1995). Size and book-to-market factors in earnings and returns. *Journal of Finance*, 50, 131–155.

Gao, B., Yang, C. (2018). Investor trading behavior and sentiment in futures markets. *Emerging Markets Finance and Trade*, 54(3), 707-720.

Guo, M., Li, Z., Tu, Z. (2012). A unique "T + 1 trading rule" in China: theory and evidence. *Journal of Banking & Finance*, 36, 575-583.

Hauser, F., Huber, J. (2012). Short-selling constraints as cause for price distortions: an experimental study. *Journal of International Money and Finance*, 31, 1279-1298.

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

Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance* 56, 1533–1597.

Hirshleifer, D., Hsu, P.-H., Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics* 107, 632–654.

Hirshleifer, D., Lim, S.S., Teoh, S.H. (2009). Driven to distraction: extraneous events and underreaction to earnings news. *Journal of Finance* 64, 2289–2325.

Hong, H., Lim, T., Stein, J.C. (2000). Bad news travels slowly: size, analyst coverage and the profitability of momentum strategies. *Journal of Finance*, 55, 265–295.

Hong, H., Stein, J.C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54, 2143–2184.

Hribar, P., McInnis, J. (2012). Investor sentiment and analysts' earnings forecast errors. *Management Science*, 58, 293-307.

Huang, D., Wen, F., Yang, X. (2009). Investor sentiment index and empirical evidence from China's stock market. *Journal of Systems Science & Mathematical Sciences*, 29(1), 1-13.

Huang, J., Su, C., Joseph, N. L., Gilder, D. (2018). Monitoring mechanisms, managerial incentives, investment distortion costs, and derivatives usage. *British Accounting Review*, 50(1), 93-141.

Huang, Y., Qiu, H., Wu, Z. (2016). Local bias in investor attention: evidence from China's internet stock message boards. *Journal of Empirical Finance*, 38, 338-354.

Hvidkjaer, S. (2008). Small Trades and the Cross-Section of Stock Returns. *Review of Financial Studies*, 21, 1123-1151.

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

Jones, C. M. (2002). A century of stock market liquidity and trading costs. Available at SSRN 313681.

Joseph, K., Wintoki, M., B., Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: evidence from online search. *International Journal of Forecasting*, 27, 1116-1127.

Kahneman, D., Tversky, A. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Kim, S.-H., Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization*, 107, 708-729.

Kumar, A. (2009). Hard-to-value stocks, behavioral biases, and informed trading. *Journal of Financial and Quantitative Analysis*, 44, 1375-1401.

Kumar, A., Lee, C. M. C. (2006). Retail Investor Sentiment and Return Comovements. *Journal of Finance*, 61(5), 2451-2486.

Lakonishok, J., Shleifer, A., Vishny, R.W. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance*, 49, 1541–1578.

Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, 105, 1-28.

Leung, H., Ton, T. (2015). The impact of internet stock message boards on cross-sectional returns of small-capitalization stocks. *Journal of Banking & Finance*, 55, 37-55.

Li, R., Li, N., Li. J, Wu, C. (2018a). Short selling, margin buying and stock return in China market. *Accounting and Finance*, 58, 477-501.

Li, X., Shen, D., Zhang, W. (2018b). Do Chinese internet stock message boards convey firm-specific information? *Pacific-Basin Finance Journal*, 49, 1-14.

Mian, G. M., Sankaraguruswamy, S. (2012). Investor sentiment and stock market response to earnings news. *Accounting Review*, 87, 1357-1384.

Ni, Z.-X., Wang, D.-Z., Xue, W.-J. (2015). Investor sentiment and its nonlinear effect on stock returns—New evidence from the Chinese stock market based on panel quantile regression model. *Economic Modelling*, 50, 266-274.

Odean, T. (1999) Do investors trade too much? *American Economic Review* 89, 1279–1298.

Penman, S., Reggiani, F. (2013). Returns to buying earnings and book value: accounting for growth and risk. *Review of Accounting Studies*, 18, 1021–1049.

Qiao, K., Dam, L. (2020).The overnight return puzzle and the "T+1" trading rule in Chinese stock markets. *Journal of Financial Markets*, 100534. https://doi.org/10.1016/j.finmar.2020.100534.

Qiu, L., & Welch, I. (2004). Investor sentiment measures. *National Bureau of Economic Research*, working paper, No. w10794.

Sabherwal, S., Sarkar, S. K., Zhang, Y. (2011). Do internet stock message boards influence trading? Evidence from heavily discussed stocks with no fundamental news. *Journal of Business Finance & Accounting*, 38, 1209-1237.

Scheinkman, J. A., Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy*, 111, 1183-1220.

Sharpe, W. F. (1970). Portfolio theory and capital markets. McGraw-Hill College.

Shen, D., Li, X., Zhang, W. (2018). Baidu news information flow and return volatility: evidence for the sequential information arrival hypothesis. *Economic Modelling*, 69, 127-133.

Siganos, A., Vagenas-Nanos, E., Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior & Organization*, 107, Part B, 730-743.

Solomon, D.H. (2012). Selective Publicity and Stock Prices. *Journal of Finance,* 67, 599–637.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., Welpe, I. M. (2014). Tweets and Trades: The Information Content of Stock Microblogs. *European Financial Management*, 20(5), 926-957.

Tumarkin, R., Whitelaw, R. F. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*, 57, 41-51.

Wysocki, P. (1998). Cheap talk on the web: the determinants of postings on stock message boards. University of Michigan Business School Working Paper (98025).

Yao, J., Ma, C., & He, W. P. (2014). Investor herding behavior of Chinese stock market. *International Review of Economics & Finance*, 29, 12-29.

Zhang, B. (2020). T+1 trading mechanism causes negative overnight return. *Economic Modelling*, 89, 55-71.

Zhang, Y., Song, W., Shen, D., Zhang, W. (2016). Market reaction to internet news: Information diffusion and price pressure. *Economic Modelling*, 56, 43-49.

Table 1. Descriptive statistics and pairwise correlations

Panel A reports the descriptive statistics for the variable decile rankings based on the average weekly ISMB postings. The observations in all variables are winsorized at ±1% before estimation. The variables are based on non-overlapping weeks. Daily stock prices are from CSMAR. The stock prices are adjusted for stock splits, and stock and cash dividends. Weekly ISMB postings for stock *i* during week *w* is the sum of the daily postings beginning on the opening of Wednesday of week *w* and ending on the closing of Tuesday of week *w*+1. The weekly total return for week *w* is the compounded daily return beginning on Wednesday of week *w* and ending on Tuesday of week *w*+1. All the stocks are ranked each week in ascending order according to their ISMB postings and then partitioned into deciles. For a given decile, the weekly ISMB postings for week *w* is the average of ISMB postings for the stocks in that decile. The average weekly total return for week *w* is the average of the weekly total returns of the stocks in each decile. Return volatility for a given week *w* is the standard deviation of daily returns during each week. Beta for week *w* is the average daily beta during the week. Firm size (capitalization) for week *w* is the average daily market capitalization values during each week. Firm age is the number of years since the firm was listed, calculated as of the Wednesday of week *w*. Profitability for week *w* is the equity return for the fiscal quarter of which the week lies. Price-to-earnings ratio is the quarterly stock price per share divided by net income per share before extraordinary items. Book-to-market (BM) ratio is the weekly book-to-market value based on the end of the fiscal quarter book value of equity divided by the average daily price per share in the week. Momentum is the cumulative return over *w*–47 to *w* –1. Stock turnover is stock turnover in week *w* based on the total daily stock trading volume during each week. Panel B reports the pairwise correlations. Beginning from the left-hand side of Panel B, the down triangular matrix shows the Pearson correlation, r_p coefficients whereas the upper triangular matrix shows the Spearman rank correlation, r_s coefficients. [a], [b] and [c] indicate statistically significant levels at 1%, 5%, 10% respectively.

Panel A: Descriptive statistics

Weekly ISMB postings by decile	Weekly ISMB postings	Weekly total return	Beta	Return volatility	Firm size	Firm age	Profitability	Price-to-earnings	Book-to-market	Momentum	Stock turnover
1（lowest）	8.03	-0.0014	1.0257	0.0152	16.3198	14.5028	0.0495	46.9976	1.5121	0.0296	0.0373
2	20.43	-0.0039	1.0440	0.0167	16.3992	14.5622	0.0457	41.3921	1.6577	0.0392	0.0388
3	31.39	-0.0033	1.0572	0.0179	16.5270	14.5664	0.0497	38.9736	1.6783	0.0649	0.0438
4	43.06	-0.0030	1.0637	0.0187	16.6075	14.8008	0.0522	38.6781	1.7000	0.0887	0.0494
5	56.74	-0.0014	1.0642	0.0202	16.7093	14.6655	0.0547	39.3132	1.6698	0.1269	0.0573
6	74.03	-0.0003	1.0665	0.0214	16.7874	14.5753	0.0561	38.6463	1.6674	0.1576	0.0642
7	98.12	0.0025	1.0620	0.0228	16.8879	14.6394	0.0599	40.4326	1.6215	0.1899	0.0745
8	135.50	0.0058	1.0593	0.0248	17.0123	14.5292	0.0617	41.3990	1.5836	0.2522	0.0880
9	205.32	0.0096	1.0588	0.0273	17.1505	14.4971	0.0644	42.5382	1.5425	0.3185	0.1053
10（highest）	445.57	0.0255	1.0704	0.0334	17.4428	13.9423	0.0672	47.1719	1.4622	0.4619	0.1488
Mean	111.85	0.0030	1.0575	0.0218	16.7915	14.5277	0.0561	41.5376	1.6093	0.1730	0.0708
Median	65.00	0.0000	1.0432	0.0179	16.7354	15.4164	0.0442	23.4474	1.0591	0.1125	0.0432
Std. dev.	135.91	0.0590	0.4098	0.0158	0.8924	5.8778	0.0753	61.1414	1.5733	0.4393	0.0790

Table 1. (Cont'd)

Panel B: Spearman rank and Pearson correlations coefficient matrix											
	ISMB postings	Total return	Beta	Return volatility	Firm size	Firm age	Profitability	Price-to-earnings	Book-to-market	Momentum	Stock turnover
ISMB postings	–	0.1032[a]	0.0598[a]	0.3172[a]	0.3607[a]	-0.0156[a]	0.0765[a]	0.0316[a]	-0.0071[b]	0.2755[a]	0.3956[a]
Total return	0.1520[a]	–	-0.0138[a]	0.0310[a]	0.0018	-0.0208[a]	0.0107[a]	0.0205[a]	-0.0311[a]	0.0130[a]	0.1255[a]
Beta	0.0588[a]	-0.0077[a]	–	0.3229[a]	0.0519[a]	-0.0111[a]	-0.0884[a]	0.1006[a]	0.1484[a]	0.0428[a]	0.3457[a]
Return volatility	0.3136[a]	0.0354[a]	0.2890[a]	–	0.1401[a]	-0.0524[a]	-0.0182[a]	0.1317[a]	-0.0471[a]	0.3118[a]	0.6245[a]
Firm size	0.1009[a]	0.0117[a]	0.4859[a]	0.3112[a]	–	-0.0276[a]	0.2726[a]	0.0261[a]	0.0326[a]	0.1543[a]	-0.0055[c]
Firm age	-0.0379[a]	-0.0220[a]	0.0046	-0.0536[a]	0.0177[a]	–	-0.0191[a]	-0.1015[a]	0.1938[a]	-0.1088[a]	0.0011
Profitability	0.0685[a]	0.0131[a]	-0.0465[a]	-0.0216[a]	0.0834[a]	-0.0334[a]	–	-0.0371[a]	-0.2128[a]	0.1096[a]	-0.0524[a]
Price-to-earnings	0.0412[a]	0.0107[a]	0.0991[a]	0.0745[a]	0.0600[a]	-0.0411[a]	-0.1114[a]	–	-0.2663[a]	0.1006[a]	0.2023[a]
Book-to-market	-0.0215[a]	-0.0286[a]	0.0950[a]	-0.0638[a]	0.1330[a]	0.1392[a]	-0.1650[a]	-0.1526[a]	–	-0.2744[a]	-0.1195[a]
Momentum	0.2875[a]	0.0107[a]	0.0254[a]	0.3667[a]	-0.0425[a]	-0.1153[a]	0.0946[a]	0.0523[a]	-0.2396[a]	–	0.4018[a]
Stock turnover	0.4208[a]	0.1725[a]	0.2278[a]	0.5668[a]	0.1894[a]	-0.0154[a]	-0.0613[a]	0.1235[a]	-0.1323[a]	0.4365[a]	–

Table 2. Short-run persistence for weekly ISMB postings and subsequent weekly returns

Panel A reports the average weekly ISMB postings for week *w* to *w* +4 for each decile. We calculate the average ISMB postings over each week for each decile. Panel B reports the average weekly total return over the next four weeks, for the stocks in each weekly ISMB decile. We rank all stocks in each week *w* of the sample period, in ascending order, according to the weekly ISMB postings and then partition the sample into deciles. [a], [b] and [c] indicate statistically significant levels at 1%, 5%, 10% respectively.

Panel A. Short-run persistence in ISMB postings

Weekly ISMB postings by decile	Average weekly ISMB postings Week *w*	Week *w* +1	Week *w* +2	Week *w* +3	Week *w* +4
1（lowest）	13.33	25.64	29.80	31.92	34.14
2	27.88	44.49	49.17	52.03	53.53
3	39.90	56.50	61.33	64.51	67.08
4	52.22	68.11	73.88	76.18	77.35
5	65.56	80.51	85.50	87.20	89.13
6	81.47	93.46	97.86	99.63	100.12
7	102.39	108.91	112.07	112.76	114.13
8	132.88	133.61	134.59	134.38	134.26
9	185.96	172.04	167.36	166.14	165.15
10（highest）	348.02	285.70	265.68	256.77	249.26
10-1	334.68[a]	260.06[a]	235.88[a]	224.85[a]	215.12[a]
t-ratio	(39.45)	(34.77)	(33.98)	(32.64)	(32.02)
Median (10)-Median (1)	319.77	253.20	231.82	218.91	207.43
Wilcoxon sign-ranks statistic	19.09[a]	19.01[a]	18.99[a]	18.97[a]	18.95[a]

Panel B. Subsequent week returns

Weekly ISMB postings by decile	Average weekly stock returns Week *w*	Week *w* +1	Week *w* +2	Week *w* +3	Week *w* +4
1（lowest）	-0.0024	0.0045[b]	0.0047[b]	0.0045[b]	0.0048[b]
	-(1.14)	(2.01)	(2.12)	(2.03)	(2.19)
2	-0.0025	0.0043[c]	0.0044[c]	0.0041[c]	0.0032
	-(1.10)	(1.79)	(1.84)	(1.76)	(1.39)
3	-0.0019	0.0046[c]	0.0040[c]	0.0037	0.0036
	-(0.78)	(1.83)	(1.70)	(1.55)	(1.52)
4	-0.0018	0.0037	0.0039	0.0037	0.0029
	-(0.78)	(1.57)	(1.60)	(1.57)	(1.23)
5	-0.0001	0.0042[c]	0.0037	0.0036	0.0028
	-(0.04)	(1.74)	(1.54)	(1.49)	(1.18)
6	0.0008	0.0046[c]	0.0040	0.0030	0.0033
	(0.33)	(1.88)	(1.64)	(1.24)	(1.36)
7	0.0019	0.0030	0.0033	0.0029	0.0033
	(0.76)	(1.26)	(1.38)	(1.19)	(1.31)
8	0.0048[c]	0.0028	0.0030	0.0026	0.0022
	(1.93)	(1.15)	(1.24)	(1.11)	(0.91)
9	0.0093[a]	0.0014	0.0016	0.0019	0.0020
	(3.78)	(0.59)	(0.66)	(0.77)	(0.78)

Table 2, Panel B (Cont'd)

Weekly ISMB postings by decile	Average weekly stock returns				
	Week *w*	Week *w* +1	Week *w* +2	Week *w* +3	Week *w* +4
10（highest）	0.0198[a]	0.0010	0.0005	0.0016	0.0008
	(7.55)	(0.38)	(0.20)	(0.62)	(0.33)
10-1	0.0222[a]	-0.0035	-0.0042	-0.0029	-0.0040
t-ratio	(6.56)	-(1.05)	-(1.24)	-(0.84)	-(1.17)
Median(10)-Median(1)	0.0203	-0.0003	-0.0019	-0.0003	-0.0023
Wilcoxon sign-ranks statistic	6.83[a]	-0.88	-1.32	-0.79	-1.35

Table 3. Short-run persistence of ISMB postings, by firm characteristic

This table reports the ISMB postings for weeks *w*+1 to *w*+4. The t-ratios are based on the difference between the highest and lowest deciles. The stock characteristics are: i) Beta (Panel A); ii) Firm size/capitalization (Panel B); iii) Book-to-market ratio (Panel C); iv) Momentum (Panel D); and, (v) Stock turnover (Panel E). The stocks are ranked each week in ascending order according to their characteristics and then partitioned into quartiles. For each quartile, the stocks are then ranked in ascending order according to the postings and portioned into deciles. [a], [b] and [c] indicate statistical significance at a 1%, 5%, 10% respectively.

Panel A. Short-run persistence of ISMB postings sorted by Beta

Beta	Weekly ISMB postings by decile	Beta	Week *w* +1	Week *w* +2	Week *w* +3	Week *w* +4
1(lowest quartile)	1	0.5853	25.9633	28.4297	30.1098	31.7889
	10	0.5678	295.2367	270.8501	256.1121	250.0237
	(10)-(1)	-0.0175	269.2734[a]	242.4204[a]	226.0023[a]	218.2348[a]
	t-ratio	(-1.48)	(34.62)	(34.23)	(33.46)	(33.19)
2	1	0.9316	27.1124	30.6237	31.8987	35.0731
	10	0.9372	307.6872	282.4096	276.5020	265.6237
	(10)-(1)	0.0056	280.5748[a]	251.7859[a]	244.6033[a]	230.5506[a]
	t-ratio	(0.67)	(35.12)	(34.26)	(32.31)	(31.31)
3	1	1.1904	26.6972	30.2045	33.2928	34.4312
	10	1.2002	313.4905	288.8100	277.8327	268.7481
	(10)-(1)	0.0098	286.7933[a]	258.6055[a]	244.5399[a]	234.3169[a]
	t-ratio	(1.12)	(30.79)	(29.71)	(27.78)	(28.20)
4 (highest quartile)	1	1.5600	25.7083	30.6148	34.4668	37.0764
	10	1.6436	357.2575	331.7503	317.2809	305.8000
	(10)-(1)	0.0836	331.5492[a]	301.1355[a]	282.8141[a]	268.7236[a]
	t-ratio	(1.42)	(37.15)	(34.31)	(33.45)	(31.42)

Panel B. Short-run persistence of ISMB postings sorted by firm size

Firm size	Weekly ISMB postings by decile	Firm size	Week *w*+1	Week *w* +2	Week *w* +3	Week *w* +4
1 (lowest quartile)	1	15.6182	18.9232	22.2816	22.9443	25.0263
	10	15.8787	174.2482	152.1351	145.6861	137.5269
	(10)-(1)	0.2605	155.3250[a]	129.8535[a]	122.7418[a]	112.5006[a]
	t-ratio	(0.87)	(28.66)	(25.36)	(24.36)	(23.68)
2	1	16.4759	27.1147	31.2968	33.5826	34.8957
	10	16.5348	224.9551	201.6842	189.5859	182.6686
	(10)-(1)	0.0589	197.8404[a]	170.3874[a]	156.0033[a]	147.7729[a]
	t-ratio	(0.24)	(31.26)	(28.54)	(26.42)	(25.37)
3	1	17.0166	41.5952	46.4846	47.6764	48.4375
	10	17.0749	305.5596	276.2200	259.9866	251.2401
	(10)-(1)	0.0583	263.9644[a]	229.7354[a]	212.3102[a]	202.8026[a]
	t-ratio	(0.24)	(33.10)	(30.62)	(28.92)	(28.54)
4 (highest quartile)	1	17.6824	32.8182	38.7670	42.7354	45.5763
	10	18.2971	445.8345	418.1826	403.4872	391.7566
	(10)-(1)	0.6147	413.0163[a]	379.4156[a]	360.7518[a]	346.1803[a]
	t-ratio	(1.61)	(40.60)	(38.21)	(36.79)	(35.40)

Table 3. (Cont'd)

Panel C. Short-run persistence of ISMB postings sorted by the book-to-market ratio

BM	Weekly ISMB postings by decile	BM ratio	Week w +1	Week w +2	Week w +3	Week w +4
1 (lowest quartile)	1	0.3302	23.0612	26.5777	30.4334	32.3699
	10	0.3045	342.9254	318.1288	305.8879	295.9302
	(10)-(1)	-0.0257	319.8642[a]	291.5511[a]	275.4545[a]	263.5603[a]
	t-ratio	-(0.11)	(38.28)	(37.20)	(35.65)	(35.52)
2	1	0.8187	28.5878	32.9495	36.2289	38.3974
	10	0.8165	312.7718	288.3221	276.5544	265.8549
	(10)-(1)	-0.0022	284.1840[a]	255.3726[a]	240.3255[a]	227.4575[a]
	t-ratio	-(0.01)	(36.10)	(33.85)	(31.72)	(31.00)
3	1	1.5529	30.5510	36.0660	37.9421	39.4538
	10	1.6104	302.1384	277.8879	268.4063	259.4104
	(10)-(1)	0.0575	271.5874[a]	241.8219[a]	230.4642[a]	219.9566[a]
	t-ratio	(0.19)	(28.75)	(26.30)	(25.18)	(24.38)
4 (highest quartile)	1	3.5217	25.0111	29.1477	30.3080	31.8223
	10	4.0594	308.3859	283.3974	272.7299	266.3673
	(10)-(1)	0.5377	283.3748[a]	254.2497[a]	242.4219[a]	234.5450[a]
	t-ratio	(0.68)	(28.87)	(27.18)	(27.03)	(26.14)

Panel D. Short-run persistence of ISMB postings sorted by price momentum

Momentum	Weekly ISMB postings by decile	Momentum	Week w+1	Week w +2	Week w +3	Week w +4
1 (lowest quartile)	1	-0.2029	25.0486	28.4390	31.1469	32.7206
	10	-0.1846	246.0163	228.1028	221.1699	215.8983
	(10)-(1)	0.0183	220.9677[a]	199.6638[a]	190.0230[a]	183.1777[a]
	t-ratio	(0.70)	(29.63)	(28.09)	(26.90)	(26.79)
2	1	0.0457	26.4783	31.3996	34.6724	35.6942
	10	0.0513	246.5911	230.6597	223.5770	216.0872
	(10)-(1)	0.0056	220.1128[a]	199.2601[a]	188.9046[a]	180.3930[a]
	t-ratio	(0.21)	(28.47)	(26.51)	(24.84)	(24.64)
3	1	0.2363	27.0527	30.9072	32.2820	35.5135
	10	0.2451	282.4497	254.4219	243.5558	236.8623
	(10)-(1)	0.0088	255.3970[a]	223.5147[a]	211.2738[a]	201.3488[a]
	t-ratio	(0.31)	(32.67)	(30.36)	(28.83)	(27.21)
4 (highest quartile)	1	0.5497	31.5933	36.7562	40.0356	41.2100
	10	0.6481	418.4505	385.6501	368.5837	358.5265
	(10)-(1)	0.0984[s]	386.8572[a]	348.8939[a]	328.5481[a]	317.3165[a]
	t-ratio	(3.30)	(39.92)	(37.28)	(35.50)	(34.93)

Table 3. (Cont'd)

Panel E. Short-run persistence of ISMB postings sorted by stock turnover						
Stock turnover	Weekly ISMB postings by decile	Stock turnover	Week *w*+1	Week *w* +2	Week *w* +3	Week *w* +4
1 (lowest quartile)	1	0.0114	21.8085	26.7547	30.1121	32.7199
	10	0.0147	260.3380	244.4998	241.7455	240.1262
	(10)-(1)	0.0033	238.5295[a]	217.7451[a]	211.6334[a]	207.4063[a]
	t-ratio	(1.48)	(30.70)	(30.12)	(28.41)	(28.78)
2	1	0.0414	26.8341	29.7347	31.2805	33.5952
	10	0.0424	255.1503	245.7796	241.2727	240.4631
	(10)-(1)	0.0010	228.3162[a]	216.0449[a]	209.9922[a]	206.8679[a]
	t-ratio	(0.34)	(28.35)	(27.06)	(27.21)	(27.66)
3	1	0.0683	28.2237	32.7032	35.4308	35.5918
	10	0.0715	291.6783	279.2212	275.0267	267.3124
	(10)-(1)	0.0032	263.4546[a]	246.5180[a]	239.5959[a]	231.7206[a]
	t-ratio	(0.75)	(31.48)	(30.15)	(29.06)	(28.26)
4 (highest quartile)	1	0.1332	33.7091	36.4011	38.1588	40.0913
	10	0.1832	384.4783	344.8575	320.3996	310.5703
	(10)-(1)	0.0500[b]	350.7692[a]	308.4564[a]	282.2408[a]	270.4790[a]
	t-ratio	(2.27)	(38.86)	(35.52)	(33.50)	(32.13)

Table 4. Short-run persistence of ISMB postings for each quartile using hard-to-value proxies

This table reports the average weekly ISMB postings for week *w*+1 of the top and bottom deciles of stocks, constructed according to average weekly ISMB postings. The proxies are: return volatility, firm size (capitalization), firm age, profitability, and price-to-earnings ratio. We first rank all the stocks each week in ascending order according to each proxy and then partitioned the associated stocks into quartiles. For each proxy's quartile, the stocks are ranked in ascending order, according to their weekly ISMB postings. They are then partitioned into deciles. [a], [b] and [c] indicate statistical significance at a 1%, 5%, 10% respectively.

Quartile of hard-to-value measure	Weekly ISMB postings by decile	Average week *w*+1 ISMB postings				
		Return volatility	Firm size	Firm age	Profitability	Price-to-earnings
1(lowest quartile)	1	22.5347	18.9232	21.8085	32.8375	23.4456
	10	253.3822	174.2482	378.1347	330.8215	312.0456
	(10)-(1)	230.8475[a]	155.3250[a]	356.3262[a]	297.9840[a]	288.6000[a]
	t-ratio	(27.80)	(28.66)	(35.99)	(32.56)	(34.99)
2	1	26.1477	27.1147	26.0824	28.1978	27.2430
	10	268.9844	224.9551	313.4464	291.5878	303.3414
	(10)-(1)	242.8367[a]	197.8404[a]	287.3640[a]	263.3900[a]	276.0984[a]
	t-ratio	(29.66)	(31.26)	(36.95)	(31.45)	(32.92)
3	1	28.1647	41.5952	27.9536	27.9499	30.7306
	10	296.6360	305.5596	296.5622	324.2174	302.6341
	(10)-(1)	268.4713[a]	263.9644[a]	268.6086[a]	296.2675[a]	271.9035[a]
	t-ratio	(32.97)	(33.10)	(38.86)	(37.19)	(28.42)
4(highest quartile)	1	31.0545	32.8182	30.5551	20.0215	27.1250
	10	393.9881	445.8345	282.9321	329.7265	343.6835
	(10)-(1)	362.9336[a]	413.0163[a]	252.3770[a]	309.7050[a]	316.5585[a]
	t-ratio	(39.69)	(40.60)	(31.60)	(39.41)	(36.61)
(4)-(1)		70.4339[a]	100.1114[a]	-10.2055[b]	-17.3690[a]	17.5012[a]
t-ratio		(13.43)	(19.30)	-(2.10)	-(3.39)	(3.91)

Table 5. Long-run cumulative return reversals for portfolios formed according to ISMB postings

This table reports the long-run cumulative buy-and-hold abnormal returns (BHARs) of portfolios formed according to ISMB postings. Stock returns are ranked each week *w*, according to their weekly ISMB postings and then partitioned into deciles. We construct three equal-weighted portfolios as follows. The first portfolio is long in the stocks in the lowest decile. The second portfolio long in stocks in the highest decile. The third portfolio is long in stocks in the highest decile and short in stocks in the lowest decile. We hold the three portfolios for each of the next week *w* to week 5, week *w* to week 10, week *w* to week 15, week *w* to week 20, … and up to week 200, starting from week *w*. This table reports the results of the cumulative buy-and-hold abnormal returns (BHARs) portfolio returns (last two rows). [a], [b] and [c] indicate statistical significance at a 1%, 5%, 10% respectively.

Weeks	Week *w*	Week *w* to 5	Week *w* to 10	Week *w* to 20	Week *w* to 30	Week *w* to 40	Week *w* to 50	Week *w* to 60	Week *w* to 70	Week *w* to 80	Week *w* to 100	Week *w* to 150	Week *w* to 200
1st decile 1 portfolio	-0.0024	0.0162[a]	0.0365[a]	0.0793[a]	0.1214[a]	0.1645[a]	0.2013[a]	0.2341[a]	0.2502[a]	0.2561[a]	0.2647[a]	0.2767[a]	0.1954[a]
t-ratio	(-1.14)	(3.13)	(4.66)	(6.65)	(7.80)	(8.62)	(9.17)	(9.78)	(10.19)	(10.39)	(10.48)	(9.52)	(7.55)
2nd decile 10 portfolio	0.0198[a]	0.0270[a]	0.0354[a]	0.0580[a]	0.0840[a]	0.1076[a]	0.1269[a]	0.1357[a]	0.1392[a]	0.1359[a]	0.1412[a]	0.1318[a]	0.1315[a]
t-ratio	(7.55)	(4.76)	(4.36)	(4.98)	(5.45)	(5.91)	(6.35)	(6.61)	(7.02)	(7.16)	(7.75)	(6.06)	(7.08)
3rd portfolio (difference)	0.0222[a]	0.0108[a]	-0.0011	-0.0213[a]	-0.0374[a]	-0.0568[a]	-0.0744[a]	-0.0985[a]	-0.1110[a]	-0.1203[a]	-0.1236[a]	-0.1449[a]	-0.0639[a]
t-ratio	(6.56)	(3.63)	(-0.26)	(-3.91)	(-5.63)	(-7.33)	(-8.25)	(-9.69)	(-10.14)	(-10.72)	(-10.56)	(11.67)	(-7.22)

Table 6. Coefficients estimated using IV-GMM

This table reports the estimated coefficient estimates for the firm characteristics in terms of their influence on the intensity of long-run return reversals, using buy-and-hold abnormal returns (BHARs) for up to week *w* to 200. The coefficients are estimated using instrumental variables-general method of moments (IV-GMM). The corresponding robust standard errors are in parentheses. The chosen IVs are those that ensure adequate model specification in terms of no over-riding restrictions, no endogeneity, and no weak instruments. The Hansen *J* statistic is used to test of over-riding restrictions. The Hayashi *C* statistic is used to test for endogeneity. The *F*-statistic is used to test for weak instruments. The partial R^2 measures the correlation between Turnover and the IVs, after partialing out other variables. [a], [b] and [c] indicate statistical significance at a 1%, 5%, 10% respectively.

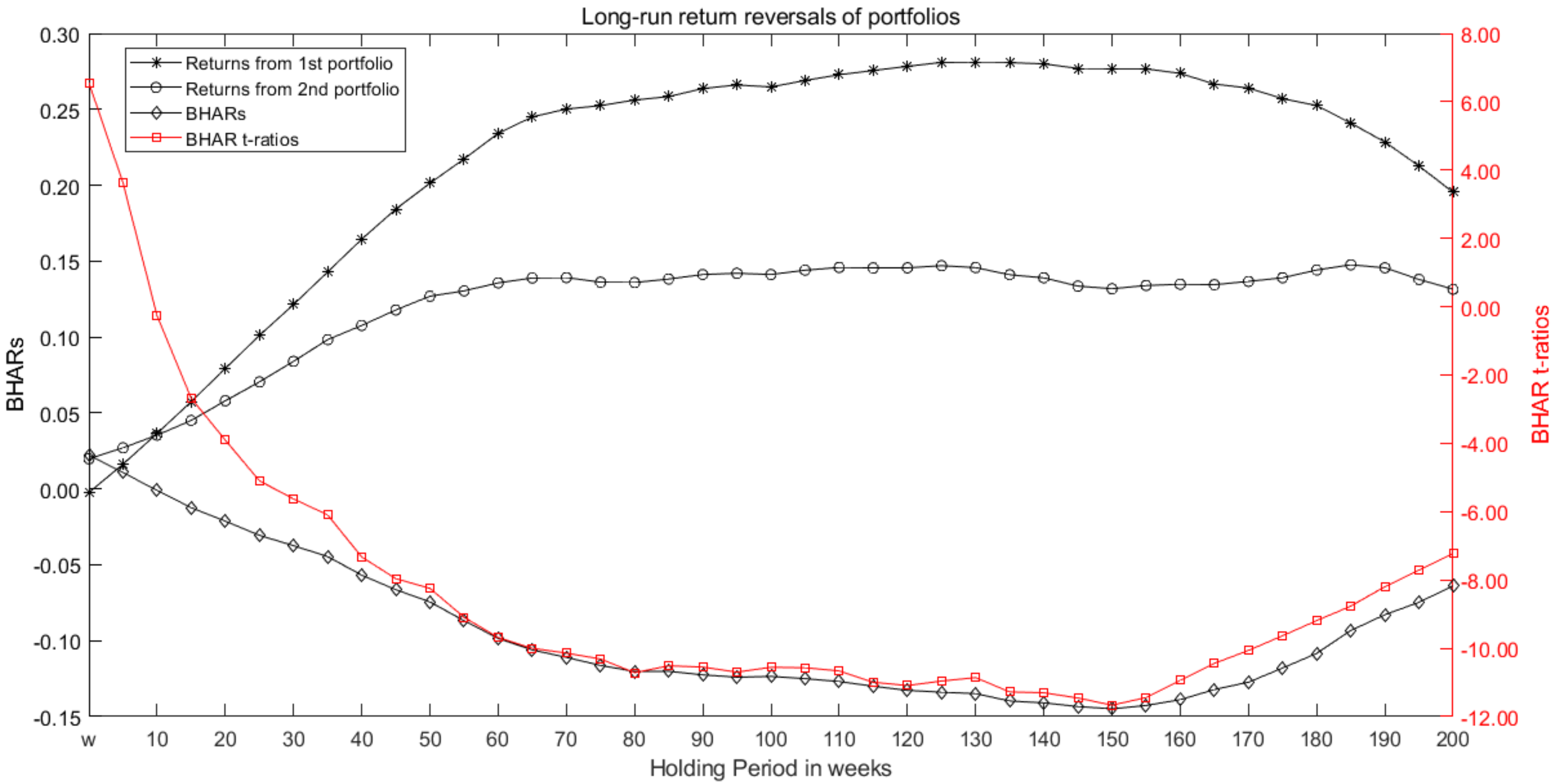
	Buy-and-hold abnormal returns (BHARs) of the third portfolio												
	Week *w*	Week *w* to 5	Week *w* to 10	Week *w* to 20	Week *w* to 30	Week *w* to 40	Week *w* to 50	Week *w* to 60	Week *w* to 70	Week *w* to 80	Week *w* to 100	Week *w* to 150	Week *w* to 200
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Stock turnover	-0.1234	3.9762[c]	5.9318[c]	2.5077	2.2524[c]	0.2084	-2.2722	-5.5132	-5.3642	-3.8168	-10.4807	-1.0648	-3.3774
	(0.3870)	(2.1628)	(3.4682)	(1.5412)	(1.3403)	(1.5961)	(2.7981)	(3.6356)	(2.6536)	(2.6939)	(23.7242)	(4.6342)	(5.2079)
Beta	-0.0708	0.4212[b]	0.6081[b]	0.4150[a]	0.3310[a]	0.3741[a]	0.4593[a]	0.4290[c]	0.4846[c]	0.5152[a]	0.2714	0.2300	1.0055[b]
	(0.0449)	(0.1829)	(0.2595)	(0.0997)	(0.0974)	(0.0842)	(0.1572)	(0.2263)	(0.2078)	(0.1524)	(0.3717)	(0.2257)	(0.4761)
Return Volatility	0.7837	-5.0235[c]	-6.7295	-3.8003[b]	-3.2680[c]	-0.9203	0.4073	2.7052	3.1878	3.1438	9.6075	0.6769	12.4847
	(0.5389)	(2.7836)	(4.0995)	(1.8373)	(1.7385)	(1.4821)	(2.0340)	(2.6071)	(1.9725)	(2.1800)	(20.4861)	(4.6499)	(15.2190)
Firm size	0.0570	-0.4297[c]	-0.6366	-0.4657[b]	-0.7196[a]	-0.4288[c]	0.0181	0.6102	0.6986	0.4138	1.7081	0.2293	0.9740
	(0.0491)	(0.2497)	(0.4035)	(0.1860)	(0.1902)	(0.2224)	(0.4295)	(0.5668)	(0.4180)	(0.4270)	(4.3823)	(0.9786)	(2.0781)
Firm age	-0.0132	0.0969[c]	0.1502[c]	0.1117[a]	0.1694[a]	0.1035[c]	0.0112	-0.1082	-0.1024	-0.0191	-0.3717	0.0457	-0.3548
	(0.0103)	(0.0538)	(0.0889)	(0.0426)	(0.0436)	(0.0539)	(0.1190)	(0.1588)	(0.1171)	(0.1200)	(1.2682)	(0.3093)	(0.2576)
Profitability	0.2557	-1.1099[c]	-1.6604[c]	-0.8079[b]	-0.5662	0.0942	0.3477	0.7649	0.7623	-0.0320	1.7863	0.3549	6.0181[a]
	(0.1462)	(0.6385)	(0.9223)	(0.3683)	(0.3686)	(0.3369)	(0.7609)	(0.9603)	(0.6819)	(0.6793)	(6.2838)	(2.3781)	(1.7572)
Price-to-earnings	-0.0079	-0.0263	-0.0551	-0.0857[a]	-0.0545[c]	-0.0675[b]	-0.0627	-0.0481	-0.0520	-0.0254	-0.0332	0.0627	0.2072
	(0.0059)	(0.0266)	(0.0431)	(0.0267)	(0.0301)	(0.0275)	(0.0417)	(0.0553)	(0.0453)	(0.0372)	(0.1355)	(0.0625)	(0.3827)
Book-to-market	-0.0015	-0.0668[c]	-0.0833	-0.1575[a]	-0.2055[a]	-0.1942[a]	-0.1702[b]	-0.1713	-0.1263	-0.1421[c]	-0.2705	0.0386	0.3208
	(0.0099)	(0.0390)	(0.0551)	(0.0308)	(0.0317)	(0.0384)	(0.0858)	(0.1209)	(0.0947)	(0.0803)	(0.4541)	(0.0579)	(0.2215)

Table 6. (Cont'd)

Buy-and-hold abnormal returns (BHARs) of the third portfolio													
	Week w	Week w to 5	Week w to 10	Week w to 20	Week w to 30	Week w to 40	Week w to 50	Week w to 60	Week w to 70	Week w to 80	Week w to 100	Week w to 150	Week w to 200
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
Momentum	-0.0310[a]	-0.1287[a]	-0.1815[b]	-0.0849[b]	-0.0369	-0.0752[b]	-0.1290[a]	-0.2067[a]	-0.2640[a]	-0.2748[a]	-0.5579	-0.2276	-0.0556
	(0.0114)	(0.0466)	(0.0774)	(0.0381)	(0.0402)	(0.0351)	(0.0493)	(0.0738)	(0.0620)	(0.0595)	(0.7844)	(0.2353)	(0.4544)
Constant	-0.6809	5.3615[c]	7.8255	5.9491[b]	9.5207[a]	5.5740[c]	-0.6068	-8.6558	-10.3670	-6.8558	-22.8343	-4.9291	-13.7992
	(0.6410)	(3.2206)	(5.1927)	(2.3986)	(2.4843)	(2.8539)	(5.3929)	(7.0947)	(5.2616)	(5.3117)	(53.7440)	(11.9099)	(32.0434)
Diagnostic tests													
No. of observations	249	244	239	229	219	209	199	189	179	169	149	99	49
Wald χ^2	24.95[a]	19.89	23.17	155.70	183.34	308.67[a]	379.68[a]	269.16[a]	401.59	499.06[a]	157.01[a]	50.56	75.27[a]
Root MSE	0.0234	0.0804	0.1211	0.0712	0.0802	0.0749	0.0905	0.1216	0.1162	0.1013	0.2149	0.0956	0.0855
Hansen J statistic (χ^2)	0.9892	1.0945	0.3843	1.4321	0.8791	9.8382[a]	0.0390	0.0361	0.5039	2.0419	0.1434	1.8650	3.3182
Hayashi C statistic (χ^2)	2.5606	3.7695[b]	1.9139	1.7328	2.1218	0.0951	0.5607	3.4420[b]	6.8473[a]	1.5414	0.9281	0.0013	1.2740
First stage Adj. R^2	0.8093	0.8065	0.8121	0.8136	0.8168	0.8123	0.8232	0.8170	0.8205	0.8154	0.7996	0.8018	0.9015
First stage partial R^2	0.0361	0.0446	0.0349	0.0239	0.0342	0.0279	0.0322	0.0323	0.0299	0.0197	0.0012	0.0083	0.0211
F-statistic for weak instrument	4.4854[a]	4.5588[a]	3.7881[b]	3.6285[b]	5.4178[a]	2.9117[b]	3.0993[b]	3.1462[b]	3.4801[b]	2.3795[c]	0.1129	0.4976	0.7049

Figure 1. Long-run buy-and-hold-abnormal returns (BHARs) using the returns from decile 1 and decile 10 portfolios

The figure shows the plots of the cumulative buy-and-hold-abnormal returns (BHARs), using the returns of the first and second portfolios, based on the decile 1 and decile 10, respectively. The figure also shows the plots of the differences between the first and the second portfolios and the corresponding t-ratios. Starting with week *w*, a buy and hold return strategy is undertaking for week *w* to week 5, week *w* to week 10, week *w* to week 15, week *w* to week 20, … and up to week 200.



Appendix I. Variable definitions

The definitions for the variables used in our study are shown below. For the weekly variables, each week *w* begins on the opening of Wednesday (00:00 on Wednesday) of week *w* and ending on the closing of Tuesday (24:00 on Tuesday) of week *w+1*. Yearly variables are considered to have 47 weeks of activity.

Variable	Definition	The studies cited below do not necessarily focus on investor sentiment and internet message boards
ISMB postings	The sum of internet stock message board postings during a week	Wysocki (1998); Li et al. (2018b)
Return volatility	The standard deviation of a stock's daily return during a week	Aboody et al. (2018)
Firm size	The average daily market capitalization of a firm during a week	Fama and French (1992)
Firm age	The number of years since the firm was listed, calculated as of the Wednesday of the week	Baker and Wurgler (2006); Aboody et al. (2018)
Profitability	A firm's profitability represented by its fiscal quarter net income divided by the average book value of equity (ROE). The profitability for a week equals the ROE of the fiscal quarter where the week lies.	Hribar and McInnis (2012); Aboody et al. (2018)
Price-to-earnings ratio	Price-to-earnings ratio is computed from the average daily price per share in the week, divided by the quarterly net income per share where the week lies	Fama and French (1992)
Beta	Beta for a week is based on CAPM. The variance and covariance measures are for the past 47 weeks.	Sharpe (1970)
Book-to-market ratio (BM)	BM ratio for a week is computed from the quarterly book value per share where the week lies, divided by the average daily price per share in the week	Fama and French (1992); Baker and Wurgler (2006)
Momentum	Return momentum for a week is the cumulative weekly return over the past year (past 47 weeks)	Baker and Wurgler (2006)
Stock turnover	Stock turnover for a week is calculated as the total trading volume in the week, divided by the average number of the outstanding shares	Tumarkin and Whitelaw (2001); Antweiler and Frank (2004)