Volume 20, Number 3 – September 2018
China Accounting and Finance Review
中国会计与财务研究
2018年9月 第20卷 第3期

How Does Foreign Sentiment Affect the Chinese Stock Markets? – Some Empirical Evidence*

Xiao Li, Dehua Shen, and Wei Zhang [1]

Received 8th of January 2018 Accepted 13th of April 2018

Abstract

In this paper, making use of a uniquely mismatched sample of investor sentiment extracted from Twitter and the Chinese stock markets (Twitter is not available in mainland China), we show that foreign sentiment has a material impact on the Chinese stock markets at both the market and industry levels. In particular, we find that (1) for the contemporaneous relationship, foreign sentiment can only influence returns on market and industry indexes when investors are in their extreme states of mind, and (2) significant sentiment contagion exists during the tranquil period, while no sentiment contagion exists during the turbulent period. Using the unique Shanghai-Hong Kong Stock Connect programme, we further prove that market purchase can propagate foreign sentiment contagion in the Chinese stock market. The results of this paper have practical implications for investors who are interested in investing in Chinese stock markets.

Keywords: Sentiment Contagion, Twitter Happiness Sentiment, Social Media, Shanghai-Hong Kong Stock Connect, Market Purchase

* This work is supported by the National Natural Science Foundation of China (71701150 and 71790594), the Young Elite Scientists Sponsorship Program of Tianjin (TJSQNTJ-2017-09) and Fundamental Research Funds for the Central Universities (63182064).

[1] Xiao Li, Assistant Professor, School of Finance, Nankai University, Tianjin, 300350, PR China; email: xiaoli@nankai.edu.cn. Dehua Shen (corresponding author), Assistant Professor, College of Management and Economics; China Center for Social Computing and Analytics, Tianjin University, Tianjin, 300072, PR China; email: dhs@tju.edu.cn. Wei Zhang, Professor, College of Management and Economics; China Center for Social Computing and Analytics, Tianjin University, Tianjin, 300072, PR China; email: weiz@tju.edu.cn.

国外投资者情绪如何影响本地市场？—来自中国股票市场的经验证据*

李晓　沈德华　张维[2]

摘要

本文通过一组独特的、不相匹配的样本：由 Twitter 文本内容所构建的投资者情绪指数和中国股票市场，研究了国外投资者情绪是如何在市场层面和行业层面影响中国股票市场。主要发现了如下的现象：（1）同期相关性研究结果表明，国外投资者情绪仅能在极端的情况下影响市场和行业指数的收益；（2）在股市平静时期，国外投资者情绪具有很明显的传染特征，而在股市波动时期，这种情绪传染消失；（3）通过沪港通所提供的自然实验环境，我们进一步证明了市场购买行为能够促进国外投资者情绪传染。上述的实证发现对于致力于投资中国股票市场的投资者有实践指导价值。

关键词：情绪传染、Twitter 情绪指数、社交媒体、沪港通、市场购买行为

* 本文受国家自然科学基金青年科学基金（71701150）、国家自然科学基金重大项目（71790594）、天津市青年人才托举工程（TJSQNTJ-2017-09）以及中央高校基本科研业务费专项资金（63182064）的资助。

2 李晓，南开大学金融学院讲师，联系地址：天津海河教育园区同砚路 38 号，邮箱：xiaoli@nankai.edu.cn。沈德华（联系作者），天津大学管理与经济学部讲师、硕士生导师，联系地址：天津市南开区卫津路 92 号，邮箱：dhs@tju.edu.cn。张维，天津大学管理与经济学部教授、博士生导师，联系地址：天津市南开区卫津路 92 号，邮箱：weiz@tju.edu.cn。

I. Introduction

The impact of investor sentiment on return predictability and market dynamics has been intensively investigated in the finance literature (Brown and Cliff, 2004; Qiu and Welch, 2006; Baker and Wurgler, 2006; Bollen *et al.*, 2011; Da *et al.*, 2015). The theoretical rationale behind these studies is that the noise trader cannot be offset by arbitrageurs and traders with erroneous beliefs (e.g. self-attribution, overreaction, and underreaction) and can actually cause changes in the autocorrelations of returns (De Long *et al.*, 1990; Barberis *et al.*, 1998; Daniel *et al.*, 1998; Hong and Stein, 1999). Recently, scholars have begun to investigate the impact of foreign sentiment on local stock market behaviour (Verma and Soydemir, 2006; Baker *et al.*, 2012; Hudson and Green, 2015; Sayim and Rahman, 2015). However, the empirical question of the relationship between foreign sentiment and Chinese stock market behaviour remains unresolved.

Our goal in this paper, therefore, is to investigate the impact of foreign sentiment on Chinese stock markets by employing two mismatched samples. In particular, to measure foreign sentiment, we consider the Twitter happiness sentiment as the proxy for foreign sentiment. This happiness sentiment is constructed by employing the natural language processing technique on 10% of all English tweets. The fact that Chinese domestic investors have no access to Twitter reduces the probability that the English tweets are mostly generated by these investors. Therefore, the mismatched samples provide us with a clean environment to investigate the impact of foreign sentiment on Chinese stock market behaviour in the sense that the proxy for foreign sentiment is the least contaminated by domestic investor sentiment. Specifically, the rationale for this investigation is that due to the ongoing financial liberalisation and international trade, the Chinese financial market is much more integrated with economically connected economies.

In particular, following the intuition that investor sentiment may have a material impact on stock markets when investors are in their extreme states of mind (Kaplanski and Levy, 2010; Lepori, 2015), we employ subgroup analysis and quantile regression analysis to investigate the impact of investor sentiment on stock returns at alternative magnitudes. Since investor sentiment is also considered as a predictor for financial variables (Baker and Wurgler, 2006; Da *et al.*, 2011), we further examine the lead-lag relationship using Granger causality analysis, and the empirical results show a significant contagious effect at both the market and industry levels. Finally, with the unique Shanghai-Hong Kong Stock Connect programme, which allows eligible institutional and individual investors from Hong Kong to trade certain stocks listed on the Shanghai Stock Exchange, we further probe the market purchase hypothesis on investor sentiment contagion (Baker *et al.*, 2012). We mainly find that market purchase can propagate sentiment contagion.

We contribute to the literature in the following three ways. First, to the best of our knowledge, we are the first to investigate the impact of foreign sentiment on Chinese stock markets. Previous studies focus exclusively on the impact of US investor sentiment on the non-US G7 countries and countries in South America (Verma and Soydemir, 2006; Bathia *et al.*, 2016). Therefore, this paper provides alternative evidence for current findings on the relationship between foreign sentiment and local stock market behaviour. We mainly find that sentiment contagion exists. Second, the Shanghai-Hong Kong Stock Connect programme, which allows investors in Hong Kong to trade shares listed on the Shanghai Stock Exchange, provides us with a natural experiment to investigate the effect of the removal of arbitrage constraints on sentiment contagion. The empirical results show that the removal of arbitrage constraints can propagate sentiment contagion. These findings are in line with the conjecture of Baker *et al.* (2012), who claim that market purchase can affect global sentiment contagion. Third, unlike the studies of Verma and Soydemir (2006), Baker *et al.* (2012), Hudson and Green (2015), and Bathia *et al.* (2016), which focus on market-level evidence of the impact of foreign sentiment, we also undertake an industry-level analysis. The empirical results show that foreign sentiment affects four industry categories: Industrials, Financials, Information Technology, and Telecommunications Services.

The remainder of this paper is organised as follows. Section II discusses the literature review from the perspective of the impact of foreign sentiment and local sentiment. Section III describes the data. Section IV provides the empirical analysis, and Section V sets forth our conclusions.

II. Literature Review

The vast majority of the empirical studies investigate the relationship between local investor sentiment and local stock market behaviour. Therefore, we first review this line of research by classifying it into three types according to the information sources from which the sentiment proxy is constructed. We then review the empirical studies on the impact of foreign sentiment.

2.1 Impact of Local Sentiment

Many studies have investigated the relationship between investor sentiment and stock returns. Owing to the choice of alternative proxies for investor sentiment, the empirical results are mixed. Generally speaking, there are three main types of proxies used in the existing literature to measure investor sentiment depending on the information sources from which the proxy is constructed. First, some studies construct investor sentiment from survey data. Using the survey on investors' intelligence by the American Association of Individual Investors (AAII), Solt and Statman (1988), Fisher and Statman (2000), Lee *et al.* (2002), and Brown and Cliff (2004) show that investor sentiment is contemporaneously related to stock returns. Lemmon and Portniaguina (2006) employ the survey on consumer sentiment from the University of Michigan and the survey on consumer confidence from The Conference Board as the proxies for investor sentiment and find that investor sentiment can forecast the returns on stocks with low institutional ownership and small stocks. On the basis of the UBS/Gallup investor sentiment survey data, Qiu and Welch (2006) find that investor sentiment has a contemporaneous correlation with stock returns.

Second, some studies use market-based variables as proxies. Lee *et al.* (1991) consider the fluctuation in discounts of closed-end funds as the proxy for individual investor sentiment and find that the sentiment affects returns on small stocks. Neal and Wheatley (1998) further show that both the discounts on closed-end funds and the net mutual fund redemptions (two proxies for individual investor sentiment) predict the size premium. Baker and Wurgler (2006) form a composite sentiment index with six market-based variables (i.e. the closed-end fund discount, the equity share in new issues, share turnover, the number and average first-day returns on IPOs, and the dividend premium) and show significant cross-sectional effects on stock returns.

Third, empiricists employ proxies derived from both the mass media and social media. Focusing on the content of the *Wall Street Journal*, Tetlock (2007) constructs a proxy for investor sentiment and finds that high pessimistic sentiment predicts high trading volume and downward pressure on stock prices followed by a reversion to fundamentals. Bollen *et al.* (2011) aroused research interest in exploring the sentiment from social media by showing that Twitter mood can increase the prediction of daily up and down changes in the closing prices of the Dow Jones Industrial Average. Since then, researchers have constructed alternative proxies for investor sentiment from other social media platforms (e.g. Facebook, Google Trends, and Google Insights), and they all find significant correlations between constructed investor sentiment proxies and market behaviour (Joseph *et al.*, 2011; Kim and Kim, 2014; Siganos *et al.*, 2014; Da *et al.*, 2015).

In addition, a few studies have investigated the impact of sentiment-altering events (e.g. aviation disasters, major sports events, and extreme weather conditions) on local stock market behaviour (Hirshleifer and Shumway, 2003; Edmans, 2007; Kaplanski and Levy, 2010). For example, Chang *et al.* (2012) find that a team's loss in the National Football League leads to lower following-day returns for locally headquartered stocks.

2.2 Impact of Foreign Sentiment

As far as we know, Verma and Soydemir (2006) were the first to investigate the impact of foreign sentiment by finding that US investor sentiment has varying degrees of impact on the stock markets of the United Kingdom, Mexico, and Brazil. Zouaoui and Nouyrigat (2011) consider the US consumer confidence index as the proxy for individual investor sentiment and find that investor sentiment increases the probability of the occurrence of market crises. Lee *et al.* (2014) discover a long-run relationship between investor sentiment in the United States and major stock markets during the subprime crisis period. By constructing one global sentiment index and six local sentiment indexes, Baker *et al.* (2012) find that the global sentiment index is a contrarian predictor of country-level returns for the stocks markets of Canada, France, Germany, Japan, the United Kingdom, and the United States. Hudson and Green (2015) and Sayim and Rahman (2015) find that US investor sentiment is highly significant in explaining UK equity returns and the Istanbul Stock Market's returns and volatility, respectively. Bathia *et al.* (2016) employ both the consumer confidence index and the market-based variables of Baker and Wurgler (2006) to proxy for investor sentiment, and they find that there exist significant spillover effects of US investor sentiment on non-US G7 stock returns.

III. Data Description

The foreign sentiment proxy is originally extracted from Twitter by the Hedonometer Team,[3] which analyses roughly 10% (50 million per day) of all English tweets dating back to September 2008. To quantify the magnitudes of the happiness sentiment, Amazon Mechanical Turk is employed to score 5,000 frequently used words from *New York Times* articles, Music Lyrics, Google Books, and Twitter messages (four corpora), resulting in a composite set of around 1,000 unique words; each of these words is scored on a nine-point scale of happiness: 1 (extremely negative), 5 (neutral), and 9 (extremely positive). For each calendar day, words contained in tweets written in English are separated, matched, and scored, generating the daily happiness sentiment. The raw happiness sentiment is obtained from http://hedonometer.org. In Figure 1, "Happiness" illustrates the evolution of the raw daily happiness sentiment. As is clearly illustrated, the raw daily happiness sentiment tends to have a high level of seasonality, as it peaks during holidays (e.g. on Thanksgiving Day, Valentine's Day, and Christmas Eve). Therefore, we employ the first-order difference of the raw daily happiness sentiment as the proxy for foreign sentiment (FS). The sample period for the foreign sentiment proxy spans from 11 September 2008 to 31 May 2017, with 3,161 (2,114) calendar (trading) days.

Since the focus of this paper is on both market-level and industry-level evidence, we choose 14 market indexes and 10 industry indexes. The selected market indexes cover the stocks traded in both the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE) as well as the stocks in different boards (i.e. the Main Board, Small and Medium Enterprise Board (SME), and ChiNext (a NASDAQ-style board)). The market indexes are as follows: SSE Composite Index (SH), SSE A-Share Composite Index (SHA), SSE B-Share Composite Index (SHB), SSE 50 Index (SH50), SSE 180 Index (SH180), SSE 380 Index (SH380), SSE and SZSE 300 Index (HS300), SZSE Composite Index (SZ), SZSE

[3] This dataset is also employed by recent studies (e.g. Zhang *et al.* (2016), Li *et al.* (2017), Zhang *et al.* (2018) and Shen *et al.* (2018)). However, those studies mainly focus on international evidence or cross-listed companies and ignore the issue of foreign sentiment.

A-Share Composite Index (SZA), SZSE B-Share Composite Index (SZB), SZSE 100 Index (SZ100), SZSE 200 Index (SZ200), SZSE Small and Medium Enterprise Index (SME), and SZSE ChiNext Index (ChiNext).

In addition, we also choose 10 categories of industry; the industry indexes are Energy Index (E), Materials Index (M), Industrials Index (I), Consumer Discretionary Index (CD), Consumer Staples Index (CS), Healthcare Index (HC), Financials Index (F), Information Technology Index (IT), Telecommunications Services Index (TS), and Utilities Index (U). In particular, we download the closing prices for the market indexes and industry indexes from the China Stock Market and Accounting Research (CSMAR) database and calculate the daily returns with the following model:

$$ret_{i,t} = log^{(CP_{i,t})} - log^{(CP_{i,t-1})}, \tag{1}$$

Table 1 Statistical Properties of the Variables

This table reports the statistical properties of the market index returns, foreign sentiment, and industry index returns. JB denotes the Jarque-Bera statistic test with the null hypothesis of Gaussian distribution. Q(20) denotes the Ljung-Box statistic test for up to the 20th order serial correlation. Happiness denotes the raw happiness sentiment, and FS denotes the foreign sentiment.

Variable	Mean	Max	Min	Median	Std.	Kurtosis	Skewness	JB	Q(20)
SH	0.0002	0.0903	-0.0887	0.0007	0.0159	7.7205	-0.6276	2102.6^{***}	76.99^{***}
SHA	0.0002	0.0903	-0.0887	0.0007	0.0159	7.7125	-0.6264	2095.3^{***}	76.92^{***}
SHB	0.0004	0.0937	-0.0989	0.0012	0.0180	9.5778	-0.7147	3993.0^{***}	74.79^{***}
SH50	0.0002	0.0900	-0.0985	0.0000	0.0174	7.1335	-0.2942	1536.2^{***}	85.63^{***}
SH180	0.0002	0.0895	-0.0946	0.0004	0.0171	7.0883	-0.4887	1557.1^{***}	85.55^{***}
SH380	0.0001	0.0743	-0.0895	0.0000	0.0155	10.1700	-1.1214	4973.6^{***}	83.33^{***}
HS300	0.0002	0.0893	-0.0915	0.0006	0.0170	6.9029	-0.5305	1441.6^{***}	78.09^{***}
SZ	0.0005	0.0852	-0.0860	0.0020	0.0182	5.7171	-0.7298	838.3^{***}	48.97^{***}
SZA	0.0005	0.0851	-0.0862	0.0020	0.0183	5.7013	-0.7283	830.1^{***}	48.80^{***}
SZB	0.0005	0.0891	-0.0826	0.0013	0.0149	7.2421	-0.5429	1689.8^{***}	64.13^{***}
SZ100	0.0003	0.0890	-0.0852	0.0008	0.0180	6.1667	-0.5781	1001.6^{***}	52.41^{***}
SZ200	0.0001	0.0636	-0.0890	0.0000	0.0148	9.9997	-1.1086	4751.0^{***}	76.09^{***}
SME	0.0007	0.0935	-0.0865	0.0022	0.0187	5.6674	-0.7316	815.7^{***}	48.13^{***}
ChiNext	0.0004	0.0676	-0.0901	0.0000	0.0185	6.2284	-0.6834	1083.1^{***}	57.01^{***}
Happiness	6.0121	6.3560	5.8700	6.0100	0.0433	9.1782	1.2464	3911.4^{***}	11958^{***}
FS	0.0025	0.1560	-0.2340	0.0020	0.0297	16.6882	-0.9144	16806.3^{***}	269.42^{***}
E	-0.0001	0.0924	-0.0973	0.0000	0.0207	6.1609	-0.2302	899.15^{***}	46.67^{***}
M	0.0000	0.0809	-0.0934	0.0005	0.0202	5.5856	-0.5119	681.52^{***}	53.42^{***}
I	0.0001	0.0955	-0.0944	0.0006	0.0191	6.6502	-0.4757	1253.96^{***}	70.35^{***}
CD	0.0005	0.0920	-0.0926	0.0008	0.0182	6.1982	-0.5780	1019.13^{***}	49.82^{***}
CS	0.0005	0.0868	-0.0954	0.0007	0.0169	6.5141	-0.3708	1136.75^{***}	80.50^{***}
HC	0.0006	0.0954	-0.1006	0.0006	0.0174	6.6915	-0.4162	1261.97^{***}	98.13^{***}
F	0.0003	0.0954	-0.1011	-0.0002	0.0190	6.8233	-0.1435	1295.40^{***}	70.40^{***}
IT	0.0003	0.0955	-0.1029	0.0012	0.0215	5.2614	-0.4572	524.36^{***}	30.85^{***}
TS	0.0003	0.0960	-0.1053	0.0006	0.0205	6.5980	-0.3380	1181.09^{***}	43.34^{***}
U	0.0001	0.0693	-0.0900	0.0003	0.0157	7.6283	-0.5692	2001.96^{***}	66.25^{***}

Table 2 Matrix of Correlation Coefficients for the Market and Industry Index Returns and Foreign Sentiment

This table reports the correlation coefficients for the market and industry index returns and foreign sentiment. Happiness denotes the raw happiness sentiment, and FS denotes the foreign sentiment. The correlation coefficients among the market and industry index returns are highly significant. The correlation coefficients between foreign sentiment and the market and industry index returns are insignificant. Foreign sentiment is positively correlated to the raw happiness sentiment.

Panel A: Market Index Returns and Foreign Sentiment

Variable	SH	SHA	SHB	SH50	SH180	SH380	HS300	SZ	SZA	SZB	SZ100	SZ200	SME	ChiNext	Happiness	FS
SH	1															
SHA	1.000	1														
SHB	0.817	0.815	1													
SH50	0.938	0.938	0.708	1												
SH180	0.980	0.980	0.786	0.980	1											
SH380	0.713	0.712	0.654	0.543	0.643	1										
HS300	0.985	0.985	0.812	0.956	0.993	0.683	1									
SZ	0.888	0.887	0.835	0.731	0.838	0.790	0.884	1								
SZA	0.887	0.887	0.833	0.730	0.837	0.789	0.883	1.000	1							
SZB	0.805	0.804	0.872	0.729	0.791	0.591	0.815	0.815	0.812	1						
SZ100	0.940	0.940	0.831	0.848	0.925	0.742	0.961	0.952	0.952	0.831	1					
SZ200	0.642	0.642	0.600	0.471	0.570	0.930	0.616	0.774	0.774	0.538	0.696	1				
SME	0.843	0.842	0.814	0.672	0.785	0.774	0.835	0.985	0.986	0.786	0.913	0.766	1			
ChiNext	0.564	0.564	0.566	0.382	0.490	0.851	0.543	0.775	0.775	0.512	0.650	0.889	0.776	1		
Happiness	0.034	0.034	0.045	0.043	0.038	0.006	0.037	0.027	0.026	0.042	0.033	-0.009	0.029	-0.003	1	
FS	0.013	0.013	0.011	0.017	0.014	0.012	0.013	0.009	0.008	0.028	0.011	-0.004	-0.001	0.006	0.360	1

Panel B: Industry Index Returns and Foreign Sentiment

Variable	E	M	I	CD	CS	HC	F	IT	TS	U	FS
E	1										
M	0.866	1									
I	0.830	0.881	1								
CD	0.774	0.845	0.884	1							
CS	0.659	0.721	0.759	0.813	1						
HC	0.625	0.713	0.763	0.813	0.797	1					
F	0.778	0.728	0.760	0.739	0.615	0.585	1				
IT	0.655	0.758	0.796	0.838	0.716	0.790	0.571	1			
TS	0.682	0.709	0.777	0.756	0.647	0.658	0.646	0.727	1		
U	0.763	0.796	0.844	0.780	0.685	0.675	0.705	0.673	0.687	1	
FS	0.025	0.016	0.005	0.010	0.009	-0.003	0.020	0.009	-0.010	0.005	1

Table 3 Results of the Unit Root Test

This table reports the results of the unit root test for all the variables. ADF denotes the augmented Dickey-Fuller test, and KPSS denotes the Kwiatkowski-Phillips-Schmidt-Shin test. The null hypothesis of the ADF test is unit root, and the null hypothesis of the KPSS test is stationarity. The optimal log length of the ADF test is chosen by the Schwarz information criterion, and the optimal bandwidth of the KPSS test is determined by the Newey-West standard errors.

Index	ADF	KPSS		ADF	KPSS
SH	-44.28***	0.0739	ChiNext	-41.86***	0.1065
SHA	-44.28***	0.0739	Happiness	-0.1495	19.5000***
SHB	-42.19***	0.0562	FS	-52.77***	0.1130
SH50	-45.42***	0.0511	E	-43.75***	0.0779
SH180	-45.02***	0.0667	M	-43.45***	0.0842
SH380	-41.55***	0.0900	I	-42.69***	0.1001
HS300	-44.66***	0.0737	CD	-44.57***	0.0832
SZ	-42.45***	0.1002	CS	-44.53***	0.0678
SZA	-42.43***	0.1007	HC	-43.44***	0.1038
SZB	-42.41***	0.0713	F	-45.93***	0.0520
SZ100	-43.93***	0.0911	IT	-43.44***	0.0560
SZ200	-41.64***	0.0960	TS	-44.42***	0.0824
SME	-42.33***	0.1050	U	-44.42***	0.0900

Notes: *** indicates significance at the 1% level.

where $CP_{i,t}$ denotes the closing prices for market index i or industry index i at day t and log denotes the natural logarithm. For the empirical analysis in the following sections, both the market index returns and the industry index returns correspond to the trading days of foreign sentiment.[4]

Table 1 reports the statistical properties for all the variables. The mean values of the 14 market return series and 10 industry return series are close to zero and the standard deviations are much larger. Except for the industry category of Energy, all the return series have positive values. The Jarque-Bera statistics (JB) reject the null hypothesis of normal distribution at the 1% significance level, indicating that all the variables are fat-tailed distributed, which is also supported by a non-zero skewness (except for Happiness) and a kurtosis larger than three. The Ljung-Box statistics (Q) reject the null hypothesis of no autocorrelations up to the 20th order for all the variables. Table 2 reports the correlation coefficients for the market and industry index returns. We find that both the market index returns and the industry index returns are highly correlated with each other and that foreign sentiment is insignificantly correlated with the market index returns and industry index returns. Table 3 reports the results of the unit root test for all the variables, including the augmented Dickey-Fuller test and the KPSS test, and we find that all the variables are stationary at the 1% significance level. Figure 1 illustrates the evolution of the market index returns and foreign sentiment, and Figure 2 shows the evolution of the industry index returns and foreign sentiment. We find that there are two distinct periods, the tranquil period

[4] Admittedly, due to the time difference, one trading day of the Chinese market could be split into two different calendar days in the United States. However, the foreign sentiment proxy employed in this paper is not exclusively generated by the Twitter users in the United States. Besides, night hours (from 1:00 a.m. to 5:00 a.m.) indicate pretty low activity on Twitter, which further alleviates the impact of the time inconsistency. We really appreciate a reviewer pointing out this issue to us.

(2008/09–2014/12) and the turbulent period (2015/01–2017/05), in Chinese stock markets. Figure 3 clearly illustrates these two periods with the SSE Composite Index and the SZSE Composite Index.

Figure 1 Market Index Returns and Foreign Sentiment

This figure illustrates the evolution of the market index returns and foreign sentiment. The start dates for the SH380, SZ200, and ChiNext are 29 November 2010, 1 September 2011, and 20 August 2010, respectively. As is clearly illustrated, large fluctuations of market returns exist after 1 January 2015.

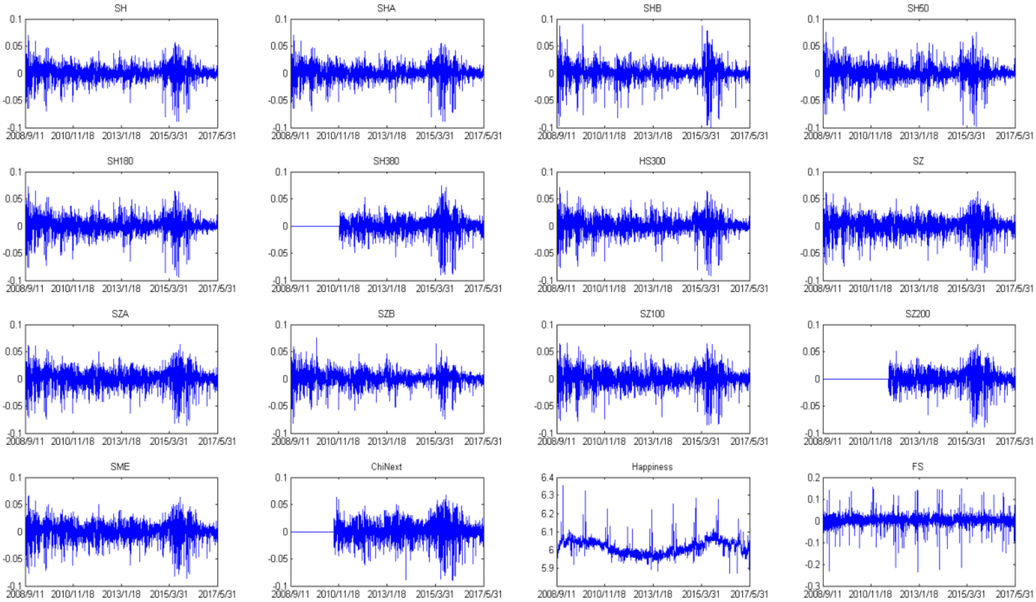


Figure 2 Industry Index Returns

This figure illustrates the evolution of the industry index returns and foreign sentiment. As is clearly illustrated, the industry categories of Healthcare, Information Technology, Financials, and Telecommunications Services have a larger range (maximum-minimum), and large fluctuations of market returns exist after 1 January 2015.

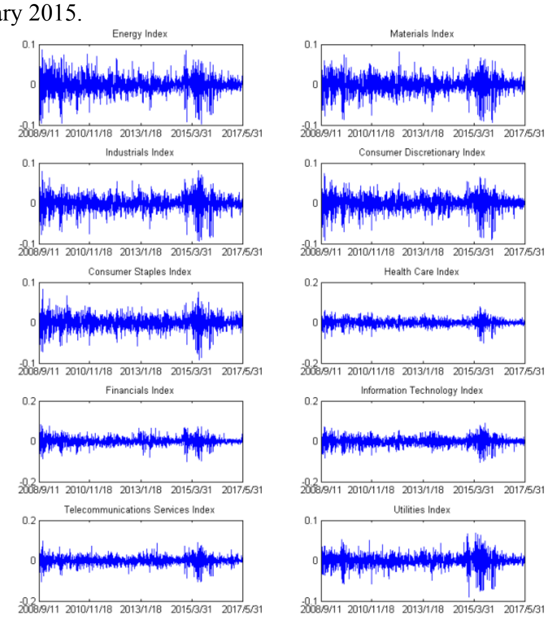
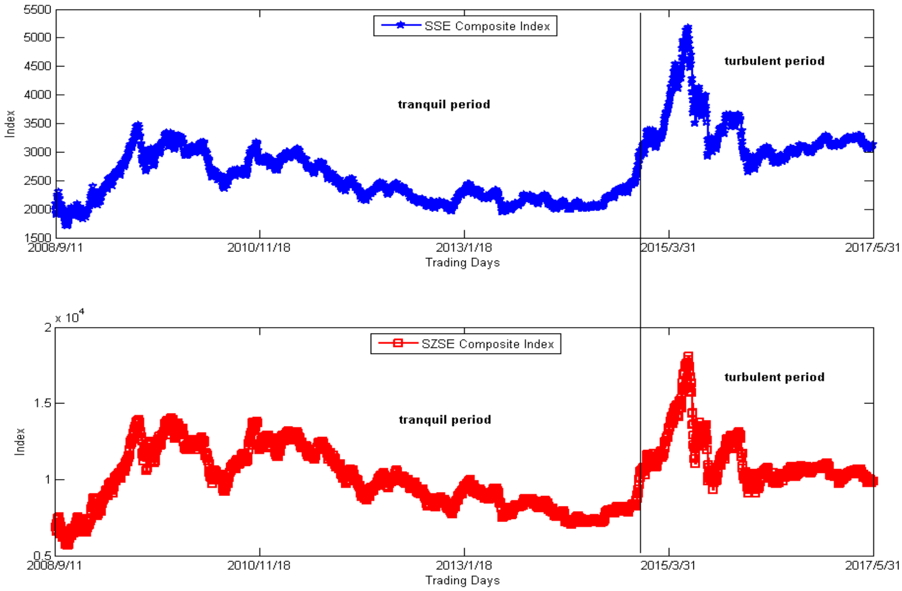


Figure 3 Tranquil and Turbulent Periods

This figure illustrates the tranquil and turbulent periods during the sample period. We roughly divide the sample period into two periods: the tranquil period (2008/09–2014/12) and the turbulent period (2015/01–2017/05).



IV. Empirical Analysis

In this section, we first employ subgroup analysis and quantile regression analysis to investigate the impact of investor sentiment on stock returns at various magnitudes. We then examine the lead-lag relationship using Granger causality. Finally, we examine the market purchase hypothesis on investor sentiment contagion by employing the Shanghai-Hong Kong Stock Connect programme, which allows eligible institutional and individual investors from Hong Kong to trade certain stocks listed on the Shanghai Stock Exchange.

4.1 Subgroup Analysis

To investigate the contemporaneous relationship between foreign sentiment and market index and industry index returns at various magnitudes, we first classify the foreign sentiment index into quintiles from the lowest to the highest, resulting in five subgroups: Lowest-FS subgroup, 2-FS subgroup, 3-FS subgroup, 4-FS subgroup, and Highest-FS subgroup. Then, we calculate the average value of the market and industry returns. Panel A of Table 4 reports the subgroup analysis for the market index returns. As shown by the results, there are significantly positive returns in the Lowest-FS subgroup and the Highest-FS subgroup. We notice that the average values do not increase monotonously from the Lowest-FS subgroup to the Highest-FS subgroup, while a significant drop is observed in the 3-FS subgroup. One possible explanation for these findings is that there exists a nonlinear relationship between investor sentiment and stock returns (Zhang *et al.*, 2018) and individual investors in Chinese stock markets are highly irrational (Feng and Seasholes, 2008; Li *et al.*, 2018). Besides, the pairwise comparisons in panel B further reveals that the return of the Highest-FS subgroup is significantly larger than that of the Lowest-, 2-, 3- and 4-FS subgroups at the 1% level. Table 5 reports the corresponding results for the relationship between foreign sentiment and industry index returns, and the results are similar to those for the market indexes. To sum up, all these findings suggest that foreign sentiment has a contemporaneous impact on Chinese stock markets which is more prominent in the Highest-FS and Lowest-FS subgroups. This is consistent with the findings of Kaplanski and Levy (2010) and Lepori (2015), which claim that investor sentiment may have a material impact on stock markets when investors are in extreme states of mind.

Table 4 Subgroup Analysis and Pairwise Comparisons of Subgroups for the Market Index Returns

Panel A reports the effect of foreign sentiment on market index returns for different subgroups. A two-sample t-test is performed to test the statistical difference from zero for each subgroup. The corresponding t-value and p-value for each subgroup are as follows: Lowest-FS subgroup (t-value=3.8448 and p-value=0.0006), 2-FS subgroup (t-value=0.4073 and p-value=0.6871), 3-FS subgroup (t-value=-5.2343 and p-value=0.0000), 4-FS subgroup (t-value=1.8441 and p-value=0.0766), and Highest-FS subgroup (t-value=5.8480 and p-value=0.0000). Values in the table are 100 times the actual market index returns. Panel B reports the pairwise comparisons of subgroups on market index returns. A two-sample t-test is performed to test the statistical differences among subgroups. As shown by the table, the return of the Highest-FS subgroup is significantly larger than that of the Lowest-, 2-, 3-, and 4-FS subgroups at the 1% level.

Panel A: Subgroup Analysis					
Index	Lowest	2	3	4	Highest
SH	0.0534	-0.0622	-0.0464	0.0044	0.1385
SHA	0.0535	-0.0628	-0.0465	0.0044	0.1385
SHB	0.0145	0.0838	-0.0502	-0.0111	0.1536
SH50	0.0541	-0.1529	-0.0370	-0.0006	0.2144
SH180	0.0641	-0.0939	-0.0416	-0.0118	0.1883
SH380	-0.0447	0.0602	-0.0167	-0.0042	0.0368
HS300	0.0654	-0.0747	-0.0320	-0.0132	0.1700
SZ	0.0857	0.0669	-0.0149	0.0362	0.0916
SZA	0.0862	0.0675	-0.0154	0.0359	0.0910
SZB	0.0162	0.0069	-0.0013	0.0676	0.1771
SZ100	0.0671	-0.0349	-0.0110	-0.0189	0.1321
SZ200	0.0092	0.0774	-0.0084	0.0044	-0.0176
SME	0.1206	0.1068	-0.0075	0.0461	0.0691
ChiNext	-0.0135	0.1485	-0.0091	0.0790	-0.0092
Mean	0.0451	0.0098	-0.0241	0.0156	0.1124
Panel B: Pairwise Comparisons of Subgroups					
Subgroup	Lowest	2	3	4	Highest
Lowest		1.3248	5.4923***	2.0423*	-2.9880***
		(0.1968)	(0.0000)	(0.0514)	(0.0061)
2			1.3884	-0.2290	-3.3409***
			(0.1768)	(0.8206)	(0.0025)
3				-4.1252***	-6.9073***
				(0.0003)	(0.0000)
4					-4.6112***
					(0.0000)

Notes: ***, **, and * indicate significance at the 1% level, 5% level, and 10% level, respectively.

Table 5 Subgroup Analysis and Pairwise Comparisons of Subgroups for the Industry Index Returns

Panel A reports the effect of foreign sentiment on market index returns for different subgroups. A two-sample t-test is performed to test the statistical difference from zero for each subgroup. The corresponding t-value and p-value for each subgroup are as follows: Lowest-FS subgroup (t-value=5.2357 and p-value=0.0000), 2-FS subgroup (t-value=-1.0562 and p-value=0.3048), 3-FS subgroup (t-value=-1.9938 and p-value=0.0615), 4-FS subgroup (t-value=0.1039 and p-value=0.9184), and Highest-FS subgroup (t-value=5.4180 and p-value=0.0000). Values in the table are 100 times the actual industry index returns. Panel B reports the pairwise comparisons of subgroups on market index returns. A two-sample t-test is performed to test the statistical differences among subgroups. As shown by the table, the return of the Highest-FS subgroup is significantly larger than that of the 2-, 3- and 4-FS subgroups at the 1% level and that of the Lowest-FS subgroup at the 10% level.

Panel A: Subgroup Analysis					
Index	Lowest	2	3	4	Highest
E	-0.0109	-0.1152	-0.0495	-0.0656	0.1791
M	0.0598	-0.0273	-0.1128	-0.0098	0.1027
I	0.0702	-0.0303	-0.0462	-0.0476	0.1248
CD	0.1250	0.0199	-0.0612	-0.0056	0.1509
CS	0.0599	0.0046	0.0366	0.0192	0.1155
HC	0.1361	0.1020	-0.0271	0.0018	0.1119
F	0.0479	-0.1725	-0.0001	-0.0105	0.2678
IT	0.0434	0.1024	-0.0614	0.0402	0.0246
TS	0.1033	-0.1849	0.0570	0.1131	0.0366
U	0.0731	-0.0330	-0.0531	-0.0191	0.0844
Mean	0.0708	-0.0334	-0.0318	0.0016	0.1198
Panel B: Pairwise Comparisons of Subgroups					
Subgroup	Lowest	2	3	4	Highest
Lowest		3.0286^{***}	4.9072^{***}	3.3496^{***}	-1.8922^{*}
		(0.0072)	(0.0001)	(0.0036)	(0.0747)
2			-0.0462	-0.9932	-3.9699^{***}
			(0.9637)	(0.3338)	(0.0008)
3				-1.4973	-5.5612^{***}
				(0.1516)	(0.0000)
4					-4.3670^{***}
					(0.0003)

Notes: *** indicates significance at the 1% level, and * indicates significance at the 10% level.

4.2 Quantile Regression Analysis

Due to the fact that investor sentiment may have a material impact on stock markets when investors are in extreme states of mind (Kaplanski and Levy, 2010; Lepori, 2015), in this section, we further investigate the contemporaneous relationship between foreign sentiment and market and industry index returns with the quantile regression (Koenker and Bassett, 1978). The quantile regression model analyses the data beyond the mean value and can be expressed as follows:

$$Index_{it} = FS_{it}^{'}\beta_\theta + \mu_{\theta it} \quad \text{with} \quad Quant_\theta(Index_{it}|FS_{it}) = FS_{it}^{'}\beta_\theta, \tag{2}$$

where $Index_{it}$ is the dependent variable denoting the market index return or industry index return, FS_{it} the independent variable denoting foreign sentiment, β the parameter to be

Table 6 Results of the Regression Analysis

This table reports the results of the quantile regression analysis for the market and industry indexes. To make comparisons, we also report the OLS results in this table. We mainly find that all the correlation coefficients are insignificant for the OLS model. However, there exist significant correlation coefficients for the market indexes (SH50, SH 180 and HS300) and industry indexes (Energy and Financials) in the lower quantiles.

	OLS	Quantile Regression			
		0.2	0.4	0.6	0.8
SH	0.0068 (0.58)	0.0232 (1.60)	0.0177 (1.78)	0.0119 (1.19)	-0.0080 (-0.51)
SHA	0.0068 (0.59)	0.0234 (1.61)	0.0178 (1.79)	0.0114 (1.13)	-0.0077 (-0.49)
SHB	0.0064 (0.48)	0.0224 (1.51)	0.0135 (1.47)	0.0064 (0.69)	-0.0144 (-1.02)
SH50	0.0102 (0.80)	0.0344 (2.56)**	0.0356 (3.43)***	0.0194 (1.80)*	0.0028 (0.15)
SH180	0.0081 (0.64)	0.0228 (1.53)	0.0314 (3.03)***	0.0175 (1.65)*	0.0018 (0.10)
SH380	0.0065 (0.57)	0.0146 (1.02)	0.0000 (0.00)	0.0030 (0.54)	0.0130 (0.77)
HS300	0.0074 (0.59)	0.0161 (1.02)	0.0238 (2.25)**	0.0128 (1.18)	0.0001 (0.01)
SZ	0.0053 (0.39)	0.0243 (1.26)	0.0203 (1.50)	0.0105 (0.80)	-0.0144 (-0.97)
SZA	0.0052 (0.39)	0.0248 (1.28)	0.0209 (1.53)	0.0094 (0.72)	-0.0147 (-0.98)
SZB	0.0140 (1.28)	0.0198 (1.45)	0.0115 (1.26)	0.0045 (0.48)	0.0069 (0.50)
SZ100	0.0065 (0.50)	0.0077 (0.43)	0.0225 (1.81)	0.0124 (0.99)	-0.0101 (-0.58)
SZ200	-0.0019 (-0.18)	0.0193 (1.42)	0.0000 (0.00)	0.0000 (0.00)	0.0063 (0.33)
SME	-0.0006 (-0.04)	0.0305 (1.52)	0.0080 (0.58)	-0.0067 (-0.49)	-0.0170 (-1.05)
ChiNext	0.0035 (0.26)	0.0109 (0.54)	0.0000 (0.00)	0.0101 (1.15)	-0.0013 (-0.06)
E	0.0174 (1.15)	0.0394 (2.10)**	0.0312 (2.20)	0.0098 (0.67)	-0.0116 (-0.62)
M	0.0106 (0.71)	0.0175 (0.93)	0.0161 (1.09)	-0.0026 (-0.17)	-0.0164 (-0.86)
I	0.0030 (0.21)	0.0090 (0.53)	0.0120 (0.92)	-0.0050 (-0.38)	-0.0163 (-0.92)
CD	0.0064 (0.48)	0.0204 (1.30)	0.0239 (1.84)	0.0123 (0.93)	-0.0208 (-1.25)
CS	0.0050 (0.40)	0.0167 (1.10)	0.0063 (0.49)	-0.0063 (-0.48)	-0.0105 (-0.67)
HC	-0.0020 (-0.16)	0.0087 (0.53)	0.0085 (0.68)	-0.0014 (-0.11)	-0.0024 (-0.16)
F	0.0130 (0.93)	0.0431 (2.80)***	0.0290 (2.62)***	0.0233 (1.98)**	0.0010 (0.05)
IT	0.0063 (0.40)	0.0249 (1.03)	0.0254 (1.58)	0.0086 (0.52)	-0.0078 (-0.40)
TS	-0.0072 (-0.48)	0.0158 (0.88)	0.0044 (0.29)	-0.0205 (-1.35)	-0.0325 (-1.82)
U	0.0026 (0.23)	0.0136 (1.09)	0.0052 (0.52)	0.0082 (0.78)	-0.0062 (-0.45)

Notes: ***, **, and * indicate significance at the 1% level, 5% level, and 10% level, respectively.

estimated, and μ the residuals. $Quant_\theta(Index_{it}|FS_{it})$ denotes the θ^{th} regression conditional quantile of $Index_{it}$ given FS_{it} and the θ^{th} $(0 < \theta < 1)$ solves the following problem:

$$\min_\beta \frac{1}{n}\left\{\sum_{i,t:Index_{it}\geq FS'_{it}\beta} \theta \left|\begin{matrix} Index_{it} \\ -FS'_{it}\beta \end{matrix}\right| + \sum_{i,t:Index_{it}\leq FS'_{it}\beta}(1-\theta)\left|\begin{matrix} Index_{it} \\ -FS'_{it}\beta \end{matrix}\right|\right\} = \min_\beta \frac{1}{n}\sum_{i=1}^{n} g_\theta\, \mu_{\theta it}, \quad (3)$$

where model (3) is solved by the linear programming method. $g_\theta(.)$ denoting the check function is defined as follows:

$$g_\theta(\mu_{\theta it}) = \begin{Bmatrix} \theta\mu_{\theta it} & if\ \theta\mu_{\theta it} < 0 \\ (\theta-1)\mu_{\theta it} & if\ \theta\mu_{\theta it} > 0 \end{Bmatrix} \quad (4)$$

In particular, we evaluate the relationship between foreign sentiment and market and industry index returns across four quantiles: 0.2, 0.4, 0.6, and 0.8. In addition, we employ the linear OLS to evaluate the mean value of foreign sentiment for the market and industry index returns. Table 6 reports the results for the OLS and quantile regressions. We mainly find that all the correlation coefficients are insignificant for the OLS model. However, there exist significant correlation coefficients for the market indexes (SH50, SH 180, and HS300) and industry indexes (Energy and Financials) in the lower quantiles. These results are consistent with our findings in section 4.1.

4.3 Lead-Lag Relationship

In this section, we further investigate the lead-lag relationship between foreign sentiment and market and industry indexes using Granger causality analysis. The Granger causality test should be analysed on stationary time series to avoid the spurious lead-lag relationship caused by the nonstationary properties. Thus, prior to the causality test, we need to employ unit root tests to analyse the nonstationary properties for the market index returns, industry index returns, and the foreign sentiment. Table 3 reports the results of the unit root test for all the variables. ADF denotes the augmented Dickey-Fuller test, and KPSS denotes the Kwiatkowski-Phillips-Schmidt-Shin test. The null hypothesis of the ADF test is unit root, and the null hypothesis of the KPSS test is stationarity. The optimal lag length of the ADF test is chosen by the Schwarz information criterion, and the optimal bandwidth of the KPSS test is determined by the Newey-West standard errors. On the basis of these two methods, we find that all the variables are stationary at the 1% significance level. Therefore, we employ the following models to illustrate the lead-lag relationship (Granger, 1988):

$$Index_t = u_{Index} + \sum_{i=1}^{p} \alpha_i Index_{t-i} + \sum_{j=1}^{q} \beta_j FS_{t-j} + \varepsilon_t \tag{5}$$

$$FS_t = u_{FS} + \sum_{i=1}^{p} \alpha_i FS_{t-i} + \sum_{j=1}^{q} \beta_j Index_{t-j} + \varepsilon_t, \tag{6}$$

where p and q denote the lag length, $Index_t$ denotes alternative market and industry index returns; FS_t denotes the foreign sentiment, α_i and β_j denote the coefficients, u_{Index} and u_{FS} denote the intercept terms, and ε_t is the regression error. The selection of the lag length (i.e. p and q) is based on the Bayesian information criterion and is determined separately for different market index returns and industry index returns.

Table 7 reports the results of the Granger causality analysis. We make four main findings. First, foreign sentiment can be a Granger-cause of the changes in B shares in both the Shanghai and Shenzhen stock exchanges (SHB and SZB). The reason is that B shares are the stocks traded in US or Hong Kong dollars by international investors and therefore changes in B shares are more subject to foreign sentiment. Second, there exists significant sentiment contagion for the largest companies on the Shanghai Stock Exchange (SH50 and SH180), in the SZSE Composite Index (SZ), and in the SZSE A-Share Composite Index (SZA). Baker *et al.* (2012) claim that market purchase and geographic approximation can induce sentiment contagion. In particular, we find that the percentages of foreign investors

Table 7 Results of the Granger Causality Analysis

This table reports the results of the Granger causality analysis between foreign sentiment and market index returns and industry index returns. The null hypothesis is no Granger causality. We employ the F-statistics and the critical value (value in parentheses) to decide the rejection. If an F-statistic is greater than the critical value, the null hypothesis of no Granger causality is rejected at a certain significance level. The symbol X→Y denotes that foreign sentiment can be a Granger-cause of the changes in market and industry index returns, and the symbol Y→X denotes the reverse direction running from market and industry index returns to foreign sentiment.

X	Y	X→Y	Y→X
Panel A: Market Index Returns			
FS	SH	3.1240 (6.6470)	4.7797 (3.8459)**
FS	SHA	3.1343 (6.6470)	4.7970 (3.8459)**
FS	SHB	7.4905 (2.4158)***	1.5904 (3.8459)
FS	SH50	7.2050 (3.3281)***	5.4566 (3.8459)**
FS	SH180	7.2426 (3.3281)***	5.3757 (3.8459)**
FS	SH380	0.8775 (6.6469)	1.4762 (3.8459)
FS	HS300	6.9084 (3.3281)***	5.1003 (3.8459)**
FS	SZ	7.2416 (3.0259)***	2.4186 (3.8459)
FS	SZA	7.1687 (3.0259)***	2.4254 (3.8459)
FS	SZB	7.8414 (3.3281)***	1.6978 (3.8459)
FS	SZ100	0.5608 (6.6470)	4.4962 (3.8459)**
FS	SZ200	1.7344 (6.6470)	1.4831 (3.8459)
FS	SME	1.7548 (6.6470)	1.9820 (3.8459)
FS	ChiNext	2.5216 (6.6470)	1.3597 (3.8459)
Panel B: Industry Index Returns			
FS	E	1.3049 (6.6469)	3.5245 (3.8459)
FS	M	0.8121 (6.6469)	5.1743 (3.8459)**
FS	I	6.1160 (3.8459)**	6.4688 (3.8459)**
FS	CD	2.0952 (6.6469)	3.3305 (6.6469)
FS	CS	2.6148 (6.6469)	1.9762 (6.6469)
FS	HC	0.8956 (6.6470)	2.8410 (6.6469)
FS	F	5.1240 (3.8459)**	5.3327 (3.8459)**
FS	IT	6.5093 (3.0259)***	1.4387 (6.6469)
FS	TS	7.6772 (3.0259)***	3.3747 (6.6469)
FS	U	1.1198 (6.6469)	4.7561 (3.8459)**

Notes: *** indicates significance at the 1% level, and ** indicates significance at the 5% level.

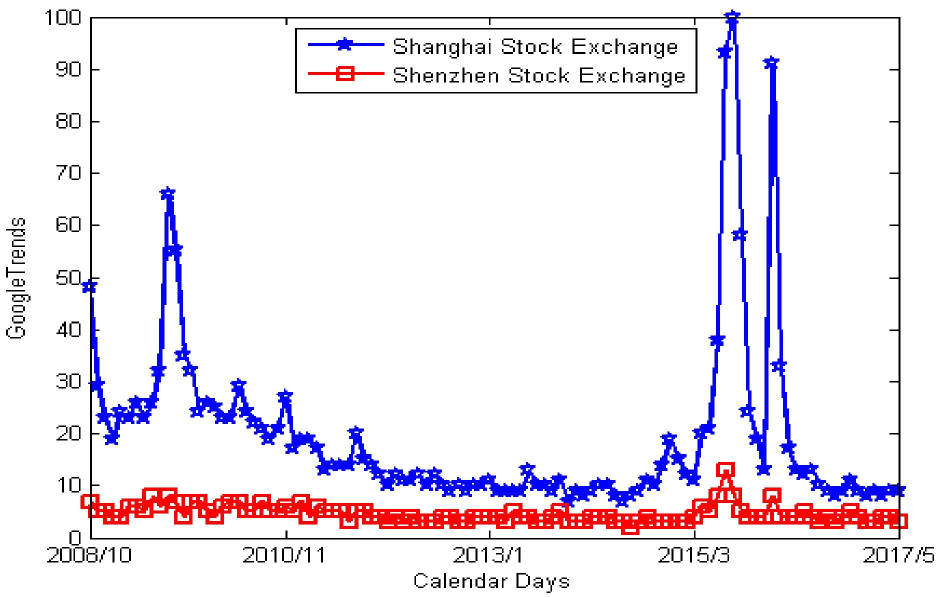
for the stocks in SH50 and SH180 are both significantly larger than those of foreign investors for other non-B-share market indexes.[5] Besides, since the Shenzhen Stock Exchange is geographically near to Hong Kong, there is significant sentiment contagion on the Shenzhen Stock Exchange overall. Third, foreign sentiment affects four industry index returns: Industrials (I), Financials (F), Information Technology (IT), and

[5] Since statistics on the holdings of foreign investors are rare, we manually collect the data from the website http://fund.jrj.com.cn/qfii/. We use the quarterly report on individual stocks' holding data to calculate the average foreign investor holding for the market indexes and then make comparisons among the market indexes.

Telecommunications Services (TS). Compared with other industry categories, these four industries are more capital intensive and require a substantial amount of investment. According to the World Investment Report 2017, China is still a major recipient of foreign direct investment (FDI). Thus, foreign sentiment propagates through these investment channels. Fourth, both the market and industry index returns can also be a Granger-cause of the changes in foreign sentiment and the effect is more pronounced for the Shanghai Stock Exchange. Figure 4 illustrates the Google Trends for the Shanghai and Shenzhen stock exchanges. As Google Trends records how often specific keywords have been searched for on Google, we find that foreign investors pay more attention to the Shanghai Stock Exchange. In that sense, foreign sentiment is more sensitive to the changes in market index returns of the Shanghai Stock Exchange.

Figure 4 Google Trends for Shanghai and Shenzhen Stock Exchanges

This figure illustrates the evolution of the Google Trends for both the Shanghai and Shenzhen stock exchanges. Google Trends records how often specific keywords have been searched for on Google. According to Da *et al.* (2011) and Zhang *et al.* (2013), Google Trends captures worldwide attention, except for in China.



4.4 Subperiod Analysis for the Lead-Lag Relationship

As is clearly illustrated in Figure 3, there are two distinct periods during the sample period: the tranquil period (2008/09–2014/12) and the turbulent period (2015/01–2017/05). Tables 8 and 9 report the results for the Granger causality analysis during the tranquil period and turbulent period, respectively. We mainly find that there exists significant sentiment contagion during the tranquil period and no sentiment contagion during the turbulent period.

The absence of a sentiment contagion effect during the turbulent period is consistent with the predictions of Cooper and Gutierrez (2004) and Chung *et al.* (2012), which claim that stock prices revert back to fundamentals and are no longer driven by sentiment.

Table 8 Results of the Granger Causality Analysis during the Tranquil Period

This table reports the Granger causality analysis for the market and industry index returns and foreign sentiment during the tranquil period. The null hypothesis is no Granger causality. We employ the F-statistics and the critical value (value in the parentheses) to decide the rejection. If an F-statistic is greater than the critical value, the null hypothesis of no Granger causality is rejected at a certain significance level. The symbol X→Y denotes that foreign sentiment can be a Granger-cause of the changes in market and industry index returns, and the symbol Y→X denotes the reverse direction running from market and industry index returns to foreign sentiment.

X	Y	X→Y	Y→X
	Panel A: Market Index Returns		
FS	SH	4.4349 (3.8476)**	7.7290 (6.6516)***
FS	SHA	4.4246 (3.8476)**	7.7479 (6.6516)***
FS	SHB	8.6796 (1.9774)***	3.1221 (2.7088)*
FS	SH50	5.8507 (3.8476)**	7.7729 (6.6516)***
FS	SH180	6.8003 (6.6516)***	7.3579 (6.6516)***
FS	SH380	4.2216 (3.8476)**	2.7661 (2.7088)*
FS	HS300	7.7540 (6.6516)***	7.2701 (6.6516)***
FS	SZ	3.3779 (2.7088)*	3.7262 (2.7088)*
FS	SZA	3.4061 (2.7088)*	3.7322 (2.7088)*
FS	SZB	8.1021 (2.5230)***	2.4531 (2.7088)
FS	SZ100	0.9516 (2.7088)	7.1841 (6.6516)***
FS	SZ200	7.5112 (6.6516)***	2.0691 (2.7088)
FS	SME	4.1594 (3.8476)**	2.6910 (2.7088)
FS	ChiNext	8.6664 (6.6515)***	1.4721 (2.7088)
	Panel B: Industry Index Returns		
FS	E	0.9415 (6.6516)	6.0130 (3.8476)**
FS	M	0.8860 (6.6516)	7.3685 (6.6516)***
FS	I	9.0664 (6.6516)	8.4192 (6.6516)***
FS	CD	2.1963 (6.6516)	4.3189 (3.8476)**
FS	CS	2.3142 (6.6516)	2.4938 (2.7088)
FS	HC	1.2548 (6.6516)	2.3972 (2.7088)
FS	F	7.5609 (2.3326)***	8.0611 (6.6516)***
FS	IT	7.9555 (3.0293)***	1.4894 (2.7088)
FS	TS	7.9370 (6.6516)***	6.7731 (6.6516)***
FS	U	6.6379 (3.0293)***	7.4250 (6.6516)***

Notes: ***, **, and * indicate significance at the 1% level, 5% level, and 10% level, respectively.

Table 9 Results of the Granger Causality Analysis during the Turbulent Period

This table reports the Granger causality analysis for the market and industry index returns and foreign sentiment during the turbulent period. The null hypothesis is no Granger causality. We employ the F-statistics and the critical value (value in the parentheses) to decide the rejection. If an F-statistic is greater than the critical value, the null hypothesis of no Granger causality is rejected at a certain significance level. The symbol X→Y denotes that foreign sentiment can be a Granger-cause of the changes in market and industry index returns, and the symbol Y→X denotes the reverse direction running from market and industry index returns to foreign sentiment.

X	Y	X→Y	Y→X
Panel A: Market Index Returns			
FS	SH	0.1441 (2.7142)	1.6798 (2.7142)
FS	SHA	0.1419 (2.7142)	1.6792 (2.7142)
FS	SHB	0.8995 (2.7142)	2.0562 (2.7142)
FS	SH50	2.1918 (2.7142)	1.7000 (2.7142)
FS	SH180	1.2555 (2.7142)	1.6613 (2.7142)
FS	SH380	0.4716 (2.7142)	1.7404 (2.7142)
FS	HS300	0.5361 (2.7142)	1.6669 (2.7142)
FS	SZ	1.1678 (2.7142)	1.7061 (2.7142)
FS	SZA	1.1654 (2.7142)	1.7059 (2.7142)
FS	SZB	1.6055 (2.7142)	1.8216 (2.7142)
FS	SZ100	0.4959 (2.7142)	1.7137 (2.7142)
FS	SZ200	1.0980 (2.7142)	1.6641 (2.7142)
FS	SME	1.2298 (2.7142)	1.6760 (2.7142)
FS	ChiNext	1.6967 (2.7142)	1.6754 (2.7142)
Panel B: Industry Index Returns			
FS	E	0.1053 (2.7142)	1.8298 (2.7142)
FS	M	0.5825 (2.7142)	1.6881 (2.7142)
FS	I	0.5687 (2.7142)	1.7540 (2.7142)
FS	CD	1.8819 (2.7142)	1.6897 (2.7142)
FS	CS	0.3259 (2.7142)	1.6617 (2.7142)
FS	HC	0.2910 (2.7142)	2.0355 (2.7142)
FS	F	2.5942 (2.7142)	1.9090 (2.7142)
FS	IT	1.6424 (2.7142)	1.6779 (2.7142)
FS	TS	0.0812 (2.7142)	2.0731 (2.7142)
FS	U	0.1180 (2.7142)	1.7509 (2.7142)

4.5 Role of Market Purchase on Sentiment Contagion

In this section, we further investigate the role of arbitrage constraints in sentiment contagion by employing a natural experiment derived from the Shanghai-Hong Kong Stock

Connect programme. Simply speaking, this programme allows eligible institutional and individual investors from Hong Kong to trade certain stocks listed on the Shanghai Stock Exchange. In the initial phase, the stocks include all the constituent stocks of the SSE180 Index and the SSE380 Index, and A+H stocks (except for the stocks that are not traded in RMB and are under risk alert). This programme officially started on 17 November 2014; therefore, we divide the sample into two periods: (1) the period before and (2) the period after the implementation of Shanghai-Hong Kong Stock Connect programme. Table 10 reports the results on the impact of arbitrage constraints on sentiment contagion. We mainly find that there is significant sentiment contagion after the implementation of the Shanghai-Hong Kong Stock Connect programme, which is consistent with prediction of Baker *et al.* (2012) claiming that market purchase can affect global sentiment contagion.

To confirm that the empirical results are not a time trend phenomenon, we also divide other market and industry indexes into two corresponding periods and observe the changes. Table 11 reports the changes in sentiment contagion around the implementation of the Shanghai-Hong Kong Stock Connect programme. We find that there is no significantly increased sentiment contagion after the implementation of the programme. Therefore, we provide evidence that market purchase induced by the Shanghai-Hong Kong Stock Connect programme propagates foreign sentiment contagion in Chinese stock markets.

Table 10 Impact of Market Purchase on Sentiment Contagion

This table report the impact of market purchase on sentiment contagion. In particular, we consider the Shanghai-Hong Kong Stock Connect programme as a shock event and investigate the changes in Granger causality results around the event. The null hypothesis is no Granger causality. We employ the F-statistics and the critical value (value in the parentheses) to decide the rejection. If an F-statistic is greater than the critical value, the null hypothesis of no Granger causality is rejected at a certain significance level. The symbol X→Y denotes that foreign sentiment can be a Granger-cause of the changes in market and industry indexes.

Panel A: Before the arbitrage constraint shock		
X	Y	X→Y
FS	SH180	0.2537 (2.8938)
FS	SH380	0.1254 (2.8938)
FS	“A+H”	1.5916 (2.8870)
Panel B: After the arbitrage constraint shock		
FS	SH180	7.9109 (3.4057)***
FS	SH380	5.0694 (3.4057)***
FS	“A+H”	8.8667 (3.4057)***

Notes: ***, **, and * indicate significance at the 1% level, 5% level, and 10% level, respectively.

Table 11 Changes in Sentiment Contagion around Shanghai-Hong Kong Stock Connect Programme

This table reports the impact of market purchase on sentiment contagion. In particular, we consider the Shanghai-Hong Kong Stock Connect programme as a shock event and investigate the changes in Granger causality results around the event. The null hypothesis is no Granger causality. We employ the F-statistics and the critical value (value in the parentheses) to decide the rejection. If an F-statistic is greater than the critical value, the null hypothesis of no Granger causality is rejected at a certain significance level. The symbol X→Y denotes that foreign sentiment can be a Granger-cause of the changes in market and industry indexes.

Panel A: Before Shanghai-Hong Kong Stock Connect			Panel B: After Shanghai-Hong Kong Stock Connect		
X	Y	X→Y	X	Y	X→Y
FS	SH	0.0045 (2.8938)	FS	SH	5.5627 (3.4959)***
FS	SHA	0.0045 (2.8938)	FS	SHA	5.5672 (3.4959)***
FS	SHB	10.6904 (3.8406)***	FS	SHB	0.3108 (2.8807)
FS	SH50	0.4280 (2.8938)	FS	SH50	10.7120 (3.8406)***
FS	HS300	0.2182 (2.8938)	FS	HS300	4.3752 (3.4959)***
FS	SZ	0.1388 (2.8938)	FS	SZ	0.9233 (2.8807)
FS	SZA	0.1315 (2.8938)	FS	SZA	0.9300 (2.8807)
FS	SZB	3.9260 (2.7763)**	FS	SZB	0.4922 (2.8807)
FS	SZ100	4.6450 (3.7653)***	FS	SZ100	1.2812 (2.8807)
FS	SZ200	5.3747 (3.7653)***	FS	SZ200	0.4128 (2.8807)
FS	SME	0.1884 (2.8938)	FS	SME	0.3864 (2.8807)
FS	ChiNext	1.8913 (2.8938)	FS	ChiNext	0.5652 (2.8807)
FS	E	0.3403 (2.8938)	FS	E	0.6770 (1.9531)
FS	M	0.7985 (2.8938)	FS	M	0.4464 1.9531)
FS	I	3.1904 (2.4205)**	FS	I	0.7699 (2.8807)
FS	CD	6.3076 (3.6305)***	FS	CD	6.7333 (3.4057)
FS	CS	0.4056 (2.8938)	FS	CS	5.2792 (3.4057)***
FS	HC	1.0159 (2.8938)	FS	HC	3.5066 (3.4057)***
FS	F	5.3259 (3.6987)***	FS	F	12.7130 (3.7910)***
FS	IT	3.0042 (2.4205)**	FS	IT	5.1174 (3.4057)***
FS	TS	5.8429 (2.3965)**	FS	TS	1.1787 (2.8807)
FS	U	1.6631	FS	U	1.0115 (2.8807)

Notes: ***, **, and *indicate significance at the 1% level, 5% level, and 10% level, respectively.

V. Conclusion

Recent studies have begun to investigate the impact of foreign sentiment on local stock market behaviour. Using mismatched samples of investor sentiment extracted from Twitter and Chinese stock markets (Twitter is not available in mainland China), in this paper, we show that foreign sentiment has a material impact on Chinese stock markets at both the market and industry levels. Specifically, for the contemporaneous relationship, we find that foreign sentiment can only influence market and industry index returns when investors are in extreme states of mind. Besides, we find that there exists significant sentiment contagion during the tranquil period and no sentiment contagion during the turbulent period. Finally, with the unique Shanghai-Hong Kong Stock Connect programme, we further prove that market purchase can propagate foreign sentiment contagion in Chinese stock markets.

The results of this paper have practical implications for both international and domestic investors. In particular, investors who are interested in investing in China should take into account the impact of foreign sentiment, and the contagious effect should be considered when constructing the asset pricing model. In addition, it would be meaningful to investigate the predictive power of foreign sentiment on stocks with certain characteristics (e.g. stocks mostly held by qualified foreign institutional investors (QFII)) and the nonlinear relationship between foreign sentiment and stock returns. We leave these issues for future work.

References

Baker, M. and Wurgler, J. (2006), 'Investor Sentiment and the Cross-Section of Stock Returns', *Journal of Finance* 61 (4): 1645-1680.

Baker, M., Wurgler, J., and Yuan, Y. (2012), 'Global, local, and contagious investor sentiment', *Journal of Financial Economics* 104 (2): 272-287.

Barberis, N., Shleifer, A., and Vishny, R. (1998), 'A model of investor sentiment', *Journal of Financial Economics* 49 (3): 307-343.

Bathia, D., Bredin, D., and Nitzsche, D. (2016), 'International Sentiment Spillovers in Equity Returns', *International Journal of Finance & Economics* 21 (4): 332-359.

Bollen, J., Mao, H., and Zeng, X. (2011), 'Twitter mood predicts the stock market', *Journal of Computational Science* 2 (1): 1-8.

Brown, G. W. and Cliff, M. T. (2004), 'Investor sentiment and the near-term stock market', *Journal of Empirical Finance* 11 (1): 1-27.

Chang, S. C., Chen, S. S., Chou, R. K., and Lin, Y. H. (2012), 'Local sports sentiment and returns of locally headquartered stocks: A firm-level analysis', *Journal of Empirical Finance* 19 (3): 309-318.

Chung, S. L., Hung, C. H., and Yeh, C. Y. (2012), 'When does investor sentiment predict stock returns?', *Journal of Empirical Finance* 19 (2): 217-240.

Cooper, M. J., Gutierrez, R. C., and Hameed, A. (2004), 'Market States and Momentum', *Journal of Finance* 59 (3): 1345-1365.

Da, Z., Engelberg, J., and Gao, P. (2011), 'In Search of Attention', *Journal of Finance* 66 (5): 1461-1499.

Da, Z., Engelberg, J., and Gao, P. (2015), 'The Sum of All FEARS Investor Sentiment and Asset Prices', *Review of Financial Studies* 28 (1): 1-32.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998), 'Investor psychology and security market under- and overreactions', *Journal of Finance* 53 (6): 1839-1885.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990), 'Noise trader risk in financial markets', *Journal of Political Economy* 98 (4): 703-738.

Edmans, A., García, D., and Norli, Ø. (2007), 'Sports Sentiment and Stock Returns', *Journal of Finance* 62 (4): 1967-1998.

Feng, L. and Seasholes, M. S. (2008), 'Individual investors and gender similarities in an emerging stock market', *Pacific-Basin Finance Journal* 16 (1-2): 44-60.

Fisher, K. L. and Statman, M. (2000), 'Investor sentiment and stock returns', *Financial Analysts Journal* 56 (Mar/Apr): 16-23.

Granger, C. W. J. (1988), 'Some recent development in a concept of causality', *Journal of Econometrics* 39 (1-2): 199-211.

Hirshleifer, D. and Shumway, T. (2003), 'Good Day Sunshine: Stock Returns and the Weather', *Journal of Finance* 58 (3): 1009-1032.

Hong, H. and Stein, J. C. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *Journal of Finance* 54 (6): 2143-2184.

Hudson, Y. and Green, C. J. (2015), 'Is investor sentiment contagious? International sentiment and UK equity returns', *Journal of Behavioral and Experimental Finance* 5: 46-59.

Joseph, K., Babajide Wintoki, M., and Zhang, Z. (2011), 'Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search', *International Journal of Forecasting* 27 (4): 1116-1127.

Kaplanski, G. and Levy, H. (2010), 'Sentiment and stock prices: The case of aviation disasters', *Journal of Financial Economics* 95 (2): 174-201.

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

Koenker, R. and Bassett, G. (1978), 'Regression Quantiles', *Econometrica* 46 (1): 33-50.

Lee, W. Y., Jiang, C. X., and Indro, D. C. (2002), 'Stock market volatility, excess returns, and the role of investor sentiment', *Journal of Banking and Finance* 26 (12): 2277-2299.

Lee, C. M. C., Shleifer, A., and Thaler, R. H. (1991), 'Investor sentiment and the closed-end fund puzzle', *Journal of Finance* 46 (1): 75-109.

Lee, Y. H., Tucker, A. L., Wang, D. K., and Pao, H. T. (2014), 'Global contagion of market sentiment during the US subprime crisis', *Global Finance Journal* 25 (1): 17-26.

Lemmon, M. and Portniaguina, E. (2006), 'Consumer Confidence and Asset Prices: Some Empirical Evidence', *Review of Financial Studies* 19 (4): 1499-1529.

Lepori, G. M. (2015), 'Investor mood and demand for stocks: Evidence from popular TV series finales', *Journal of Economic Psychology* 48: 33-47.

Li, X., Shen, D., Xue, M., and Zhang, W. (2017), 'Daily happiness and stock returns: The case of Chinese company listed in the United States', *Economic Modelling* 64: 496-501.

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

Neal, R. and Wheatley, S. M. (1998), 'Do measures of investor sentiment predict returns?', *Journal of Financial and Quantitative Analysis* 33 (4): 523-547.

Pontiff, J. (1996), 'Costly Arbitrage: Evidence from Closed-End Funds', *Quarterly Journal of Economics* 111 (4): 1135-1151.

Pontiff, J. (2006), 'Costly arbitrage and the myth of idiosyncratic risk', *Journal of Accounting and Economics* 42 (1): 35-52.

Qiu, L. and Welch, I. (2006), 'Investor Sentiment Measures', Working Paper, National Bureau of Economic Research.

Sayim, M. and Rahman, H. (2015), 'An examination of U.S. institutional and individual investor sentiment effect on the Turkish stock market', *Global Finance Journal* 26: 1-17.

Shen, D., Liu, L., and Zhang, Y. (2018), 'Quantifying the cross-sectional relationship between online sentiment and the skewness of stock returns', *Physica A: Statistical Mechanics and its Applications* 490: 928-934.

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

Solt, M. E. and Statman, M. (1988), 'How Useful Is the Sentiment Index?', *Financial Analysts Journal* 44 (5): 45-55.

Tetlock, P. C. (2007), 'Giving content to investor sentiment: The role of media in the stock market', *Journal of Finance* 62 (3): 1139-1168.

Verma, R. and Soydemir, G. (2006), 'The impact of U.S. individual and institutional investor sentiment on foreign stock markets', *Journal of Behavioral Finance* 7 (3): 128-144.

Zhang, W., Li, X., Shen, D., and Teglio, A. (2016), 'Daily happiness and stock returns: Some international evidence', *Physica A: Statistical Mechanics and its Applications* 460: 201-209.

Zhang, W., Shen, D., Zhang, Y., and Xiong, X. (2013), 'Open source information, investor attention, and asset pricing', *Economic Modelling* 33: 613-619.

Zhang, W., Wang, P., Li, X., and Shen, D. (2018), 'Twitter's daily happiness sentiment and international stock returns: Evidence from linear and nonlinear causality tests', *Journal of Behavioral and Experimental Finance* 18: 50-53.

Zouaoui, M., Nouyrigat, G., and Beer, F. (2011), 'How Does Investor Sentiment Affect Stock Market Crises? Evidence from Panel Data', *Financial Review* 46 (4): 723-747.