Lingnan University

Digital Commons @ Lingnan University

Theses & Dissertations

Department of Finance and Insurance

9-22-2017

Geographic distance and the impact of investor sentiment on stock prices

Yan YANG

Follow this and additional works at: https://commons.ln.edu.hk/fin_etd

Part of the [Finance and Financial Management Commons](#)

Recommended Citation

Yang, Y. (2017). Geographic distance and the impact of investor sentiment on stock prices (Master's thesis, Lingnan University, Hong Kong). Retrieved from http://commons.ln.edu.hk/fin_etd/16/

This Thesis is brought to you for free and open access by the Department of Finance and Insurance at Digital Commons @ Lingnan University. It has been accepted for inclusion in Theses & Dissertations by an authorized administrator of Digital Commons @ Lingnan University.

Terms of Use

The copyright of this thesis is owned by its author. Any reproduction, adaptation, distribution or dissemination of this thesis without express authorization is strictly prohibited.

All rights reserved.

GEOGRAPHIC DISTANCE AND
THE IMPACT OF INVESTOR SENTIMENT
ON STOCK PRICES

YANG YAN

MPHIL

LINGNAN UNIVERSITY

2017

GEOGRAPHIC DISTANCE AND THE IMPACT OF INVESTOR
SENTIMENT ON STOCK PRICES

by
YANG Yan
杨燕

A thesis
submitted in partial fulfillment
of the requirements for the Degree of
Master of Philosophy in Business

Lingnan University

2017

ABSTRACT

Geographic Distance and the Impact of Investor Sentiment on Stock Prices

by

YANG Yan

Master of Philosophy

Based on China's stock market, this study investigates how firms' geographic distance from a financial center affects the sensitivity of stock prices to investor sentiment. I find that firms located closer to a financial center are more affected by investor sentiment than firms located far from a financial center. This distance effect holds for different geographic cutting boundaries and after excluding firms located in financial centers. Besides, using China's High Speed Railway (HSR) as an exogenous shock, I find that HSR connection significantly decreases the effect of geographic distance on the sentiment-driven stock price relationship. In addition, firms with shorter travel times to financial centers are more affected by investor sentiment than firms with longer travel times. Moreover, firms located in provinces with a high stock market participation rate are more affected by investor sentiment than other firms. And Analysts increase the frequency of favorable recommendations for firms that are located closer to financial centers when investor sentiment is high. Furthermore, firms located closer to a financial center do not have higher institutional ownership than other firms. Last but not least, firms located in more economically developed provinces are not more affected by investor sentiment than firms located in less developed provinces. Overall, my findings highlight the importance of geographic distance in explaining the effects of investor sentiment on stock prices.

DECLARATION

I declare that this is an original work based primarily on my own research, and I warrant that all citations of previous research, published or unpublished, have been duly acknowledged.

YANG Yan

Date

CERTIFICATE OF APPROVAL OF THESIS

GEOGRAPHIC DISTANCE AND THE IMPACT OF INVESTOR SENTIMENT ON
STOCK PRICES

by
YANG Yan

Master of Philosophy

Panel of Examiners :

 _____ (Prof. LO Wai Yee Agnes)	(Chairman)
 _____ (Prof. ZOU Hong)	(External Member)
 _____ (Prof. WONG Man Lai Sonia)	(Internal Member)
 _____ (Prof. GAO Jin)	(Internal Member)

Chief Supervisor :

Prof. GAO Jin

Approved for the Senate :

_____
(Prof. MOK Ka Ho Joshua)
Chairman, Postgraduate Studies Committee

22 SEP 2017

_____
Date

TABLE OF CONTENT

Chapter 1. Introduction	1
Chapter 2. Research Background	
2.1. Investor sentiment and asset valuation	15
2.2. Firm location and asset pricing.....	18
2.3. Geographic distance and information asymmetry	19
2.4. Geographic distance, familiarity and social interaction	20
2.5. Investor sentiment and sell-side analysts	22
2.6. Institutional background and geographic distribution in China	24
Chapter 3. Hypothesis Development and Empirical Results	
3.1. Proximity and sentiment effect: more information or more behavioral bias.....	27
3.2. Construct and validate the investor sentiment index in Chinese stock market	30
3.2.1. Construct the investor sentiment index in Chinese stock market	32
3.2.2. Validate the investor sentiment index in Chinese stock market.....	35
3.3 Empirical results	
3.3.1. Sample description and data source	37
3.3.2. Investor sentiment effect and distance	37
3.3.3. Chinese High Speed Railway reform as an exogenous shock.....	42
3.3.4. Additional analysis	
3.3.4.1. Stock market participation and investor sentiment effect.....	49
3.3.4.2. Analyst recommendation number and investor sentiment	51
3.3.4.3 Institutional ownership and investor sentiment	58
3.3.4.4 Gross Domestic Product and investor sentiment effect	59
Chapter 4. Conclusion, Limitations and Avenues for Future Research	60
References	94
Appendix	101

LIST OF TABLES AND FIGURES

Table 1. Summary statistics of investor sentiment data and control variables.....	62
Table 2. Validation of the sentiment index and its relation with future stock returns.....	66
Figure 1. Comparison of sentiment index	67
Table 3. Sample selection.....	68
Table 4. Number of firms in each geographic portfolio	69
Table 5. Geographic distance and investor sentiment effect.....	70
Table 6. Geographic distance and investor sentiment effect (excluding firms located in financial center).....	72
Table 7. Geographic distance and investor sentiment effect for local firms and nonlocal firms connected to HSR lines.....	74
Table 8. Geographic distance and investor sentiment effect for local firms and nonlocal firms unconnected to HSR lines.....	76
Table 9. Geographic distance and investor sentiment effect for local firms and nonlocal firms connected to HSR lines for subsamples.....	78
Table 10. Geographic distance and investor sentiment effect for local firms and nonlocal firms unconnected to HSR lines for subsamples.....	80
Figure 2. Sentiment coefficients over time before and after HSR reform	82
Table 11. Travel time and investor sentiment effect.....	83
Table 12. Stock market participation and investor sentiment effect.....	84
Table 13. Summary statistics of stock recommendations.....	86
Table 14. Number of analyst recommendations and investor sentiment effect	87
Table 15. A comparison of institutional ownership.....	92
Table 16. Gross Domestic Product and investor sentiment effect.....	93

ACKNOWLEDGEMENT

First of all, I would like to express my sincere gratitude to Prof. Michael Firth, who guided me to the academia and gave me courage in time of difficulties in the course of my research. He managed to come to my topic defense seminar even if he was suffered from cancer at that time. Soldiers pass away in the battlefield whereas scholars end up on the desk. His scholarly spirit never dies and will inhabit in my mind during my life.

Besides, I dedicate my genuine thanks to my advisor Prof. Gao Jin for the continuous support of my Mphil study and life, for his patience and immense knowledge and for his encouragements when I doubt myself. His guidance helped me in all the time of my research and writing of this thesis.

In addition to my advisor, I would like to thank the rest of my thesis committee: Prof. LO Wai Yee, Prof. Zou Hong, and Prof. Wong Man-lai, for their insightful comments, which incited me to widen my research from various perspectives.

Furthermore, I wish to extend my appreciation to Prof. Li Jingyuan, Prof. Winnie Poon, Prof. Zhao Xiaofeng and seminar participants at two Lingnan University seminars for helpful comments on the thesis, and Ms. Clara Hui, Ms. Agnes Tsui, for their kind help during my studies.

Last but not the least, I would like to thank my families and friends, especially my grandfather, YANG Mingcai, for his valuable life experience working as a teacher and my friend LI Xuan, for his academic supports to me in writing this thesis.

Chapter 1. Introduction

The impact of investor sentiment on stock prices (hereafter the investor sentiment effect) has received increasing attention from scholars during the past decade. An increasing number of studies have documented that investor sentiment can affect stock prices, and periods of high sentiment should be followed by low future stock returns as the price reverts to its fundamental value (e.g., Brown and Cliff, 2005; Tetlock, 2007; Seybert and Yang, 2012; Da et al., 2015). Furthermore, some studies find cross-sectional variations in the investor sentiment effect, with a stronger effect for firms with greater subjectivity and that are difficult to value. Specifically, Baker and Wurgler (2006, 2007) find that the investor sentiment effect is stronger for small, young, high-volatility, extreme-growth, distressed, unprofitable and non-dividend-paying stocks. Firth et al. (2015) argue that low corporate transparency limits the availability of objective firm-specific information to investors and consequently forces them to rely more on their subjective judgments in valuing corporate assets. Consistent with their expectation, they find that firms with low corporate transparency are more affected by investor sentiment than firms with high corporate transparency. This study adds to this important emerging literature by examining whether investor sentiment effects have a geographic component. Specifically, I examine whether the distance of a firm from financial centers (geographic proximity hereafter) influences the investor sentiment effect.

The basic motivation of my study comes from the literature demonstrating that the

location of firms matters for equity pricing (e.g., Coval and Moskowitz, 2001; Hong et al., 2008; El Ghouli et al., 2013). For example, Coval and Moskowitz (2001) show that local investor ownership of a firm is positively related to its future expected returns. Pirinsky and Wang (2006) find that firms headquartered in the same geographic area exhibit a strong degree of co-movement in stock returns and argue that the price formation in stock markets has a significant regional component. Two major explanations have been proposed to explain the geographic component of stock price formation. The first is the information perspective, which suggests that information regarding the value of corporate assets is regionally distributed and stock prices exhibit a regional pattern due to the incorporation of the regionally distributed information (e.g., Ivković and Weisbenner, 2005; Anand et al., 2011; Kubick and Lockhart, 2016). The second explanation, based on cognitive bias, suggests that market participants feel more familiar with local firms than nonlocal firms, and such a divergence in familiarity may induce a regional pattern of stock prices (e.g., Huberman, 2001; Korniotis and Kumar, 2013).

Building on previous research on investor sentiment and the geographic component of asset pricing, I investigate whether the location of a firm is also a determinant of the investor sentiment effect. Specifically, I examine whether the geographic distance of a firm from its financial center affects the sensitivity of stock prices to market-wide investor sentiment. I focus on the distance from a financial center because a financial center has a high concentration of financial institutions (including security firms, investment banks and mutual funds) and a high density of individual investors, which have important

implications for the regional distribution of information (e.g., Loughran and Schultz, 2005; Loughran, 2007). Furthermore, the high concentration of financial market participants can facilitate social interactions between them, which may help to spread or intensify their cognitive biases. As I explain shortly, the effect of a firm's proximity to a financial center on the investor sentiment effect is theoretically unclear and is therefore an important and interesting empirical issue for investigation.

Specifically, I propose two competing hypotheses on the impact of geographic proximity on the investor sentiment effect. The first is the information perspective, which suggests that market participants tend to have more firm-specific information regarding firms located near a financial center than distant firms. First, investors can gather more hard information—information that is quantitative, easy to store and transmit in impersonal ways, and with content that is independent of the collection process (Petersen, 2004)—about firms located near a financial center than distant firms because sell-side analysts work for brokerage firms that are predominantly located near financial centers. As a result, analysts can issue more research reports on proximate firms than on distant firms. O'Brien and Tan (2015) offer evidence that analysts are 80% more likely to cover local firms than nonlocal ones. In addition, Cheng et al. (2016) find that analysts who conduct corporate site visits can acquire more firm-specific information and thus have greater forecast accuracy than other analysts. Second, market participants can also obtain more soft information—information that is difficult to completely summarize in a numeric score (Petersen, 2004)—on proximate firms than remote firms. For instance, they may directly

talk to the managers, suppliers and employees or directly observe the daily operations of local firms. They may also have personal ties with local firms' executives. Furthermore, firms' investment decisions may be influenced by the local media, of which local market participants are also more likely to be the audience (Coval and Moskowitz, 1999; El Ghoul et al., 2013). To the extent that there is more accurate firm-specific soft and hard information on local firms than nonlocal firms, investors have more reliable information to draw on when investing in local firms. The availability of information is expected to reduce investors' subjectivity in valuing local firms and reduce their reliance on investor sentiment.

However, previous research documents that investors may tend to invest locally because they feel more comfortable investing in firms with which they are familiar but do not necessarily have more information about (the cognitive bias perspective) (e.g., Grinblatt and Keloharju 2001; Huberman 2001; Seasholes and Zhu 2010). As financial centers have a high concentration of financial institutions and individual investors and tend to be more familiar with local than nonlocal firms, firms that are closer to financial centers are more likely to be affected by familiarity-driven behavior. Furthermore, a high concentration of market participants also facilitates social interactions, which can facilitate exchanges of ideas and beliefs (Shiller et al., 1984). Previous research also shows that these social interactions can have important implications for investor behavior. For example, Hong et al. (2004) find that social interactions increase stock market participation. Han and Hirshleifer (2016) model how the bias toward sharing positive information induces more

trading. To the extent that market participants are more (less) likely to talk about firms they are familiar with when investor sentiment is high (low), I expect the perceived familiarity of local firms to increase (decrease) when investor sentiment is high (low). As a result, investors are more (less) likely to buy local firms than nonlocal firms when market sentiment is high (low), resulting in a stronger investor sentiment effect for local firms than nonlocal firms.

To summarize, the information perspective predicts that the investor sentiment effect is weaker for local than for nonlocal firms, while the behavior argument proposes that the investor sentiment effect is stronger for local than for nonlocal firms.

In this study, I aim to distinguish these two competing hypotheses using China's stock market as my testing ground. China is the world's largest emerging market, ranking second in terms of gross domestic product (GDP)¹ and third in terms of geographic area.² China's stock market is dominated by individual investors who lack experience with market downturns (Ng and Wu, 2006), which makes them more prone to optimism. The opportunities for short-selling and margin trading are very limited in China, which deters arbitrageurs from correcting for mispricing (Sharif et al., 2014). The level of corporate transparency is also low in China, which limits the possibility for investors to conduct objective evaluations of corporate assets and consequently intensifies the investor sentiment effect (Firth et al., 2015). As demonstrated by Firth et al. (2015), the investor

---

¹ Second to the United States from 2010 and afterwards, according to the IMF.

² 9,596,961 km^2 , after Russia and Canada.

sentiment effect in China tends to be stronger (in terms of magnitude and duration) than those documented in mature markets. The presence of a significant investor sentiment effect in China makes it a particularly useful venue for investigating the determinants of the effect.

China is a vast country with three well-known financial centers in Shanghai, Shenzhen and Beijing. China's two stock exchanges are located in Shanghai and Shenzhen. Beijing, as the Chinese capital, has become the financial services regulatory center of China. It is also the headquarters of the financial regulatory authorities and the four national banks. As I show in the next section, these three financial centers in China have a high concentration of financial institutions and a vast investor base, making them a potential origin of information and a venue for social interactions.

All initial public offerings (IPOs) in China must be approved by the China Securities Regulatory Commission (CSRC) (Yang, 2013). The CSRC allocates the security licenses for firms regarding IPOs and while accounting for regional development; this creates a network of geographically distributed firms across the country and thus provides huge variations in geographic distance from firm headquarters to financial centers.

More importantly, the introduction of the High Speed Railway (HSR) in China, the longest HSR system in the world, provides a valuable testing ground to conduct a

portfolio-based difference-in-difference analysis on whether the distance from financial centers really matters for the investor sentiment effect. The first HSR train was introduced on April 18, 2007.³ New lines have gradually been built across the country and the railway has become immensely popular, with an annual ridership of over 1.44 billion in 2016.⁴ An HSR connection significantly reduces the travel time required for cross-city travel and has reduced the effective distance between connected cities and financial centers. Most importantly, the placement of HSR lines should be based on comprehensive consideration of the economic development, population and resource distribution, national security, environmental concerns and social stability of each region,⁵ which have nothing to do with the general market investor sentiment. Furthermore, HSR lines were introduced in different regions at different times, thus providing the valuable cross-section variations in HSR connections required for a difference-in-difference analysis (firms connected to HSR and firms unconnected to HSR during our sample period).

My investigation proceeds in two major steps. The first step is to construct the measure of investor sentiment and validate it. I follow the approach of Baker and Wurgler (2006) and Firth et al. (2015) in constructing the measure of market-level investor sentiment with seven variables: closed-end fund discounts, market turnover, the number of IPOs, the IPO first-day returns, the share of equity issues in new financing, the growth of investment

---

³ A trial line with a speed of 200 km/h opened in 2002 between Qinhuangdao and Shenyang and was upgraded to 250 km/h in 2007. According to China's Rail Safety Management Regulations initiated in 2014, HSR refers to railway lines running at an average speed of 250 km/h or more, or passenger-dedicated intercity lines running at an initial speed of 200 km/h or more.

⁴ China Railway Yearbook 2016.

⁵ Mid-to-Long Railway Plan, 2008, by the Ministry of Railway in China.

accounts and the growth of savings deposits. While following the same methods to construct the investor sentiment measures as used by Firth et al. (2015), I extend the sample period to the end of 2015. I also follow Firth et al. (2015) in validating the investor sentiment index in China's stock market and find similar results. Specifically, I find that the investor sentiment index is negatively related to subsequent market returns from 5 months onward, and this relation becomes significant at 11 months after, which is consistent with the investor sentiment effect in China's stock market.

I then compute the geographic distance between the headquarters of a firm and its relevant financial center. As mentioned, there are three financial centers in China— Shanghai, Shenzhen and Beijing. Specifically, firms that are headquartered within a certain radius of Beijing, Shanghai or Shenzhen are considered as local firms, whereas firms that are headquartered outside the radius are considered nonlocal firms. Following previous studies (e.g., Grinblatt and Keloharju, 2001; Ivković and Weisbenner, 2005; Kang and Kim, 2008; Seasholes and Zhu, 2010), I consider radiuses of 250 miles, 100 miles and 100 kilometers. To better investigate the difference in the investor sentiment effect between local and nonlocal firms, I also compare the effect on firms located within a radius of 100 kilometers and firms outside a radius of 1,000 kilometers. As the investor sentiment effect becomes significant at 11 months and after, I focus on the horizons of 9, 12 and 15 months.

The second step is to conduct a regression analysis to discern the investor sentiment

effects. My investigation period ranges from January 2003 to December 2015 due to the lack of monthly trading accounts data before 2003. My investor sentiment composite index consists of 156 monthly observations, and the initial sample contains all nonfinancial firms listed on either the Shanghai or Shenzhen A-share stock exchanges. To examine the impact of investor sentiment on stock prices, I follow previous research in regressing future cumulative returns on sentiment and include control variables for each set of portfolios sorted by local versus nonlocal firms. I find consistent and significant evidence that local firms are more affected by investor sentiment than nonlocal firms, regardless of the criteria used to classify them as local versus nonlocal. The result is consistent with the cognitive bias perspective. This distance effect still holds after I eliminate firms headquartered right in the financial centers, suggesting that the results are driven not by financial centers but by the important role that geographic distance plays in the sentiment-driven stock price relationship. Furthermore, using China's HSR as an exogenous shock, I find that the impact of investor sentiment on stock prices varies not only between local and nonlocal firms, but also in terms of HSR connection. Specifically, I divide firms into HSR-connected and HSR-unconnected groups. The distance effect holds in the circumstance that nonlocal firms stay unconnected to HSR during the sample period. In contrast, I observe the distance effect for the HSR-connected group before the connection, but it disappears once the HSR is in operation, which provides direct evidence that upgrading the transport infrastructure can mitigate the impact of geographic distance on the investor sentiment effect. Moreover, in order to mitigate the effect of aviation, I

restrict firms located within a radius of 1500 kilometers⁶ and conduct the difference-in-difference analysis again to clarify the effect of HSR lines. The results are still holds. Furthermore, I divide firms into local and nonlocal firms according to the travel times (including railway, airline and toll road) and find that firms with less travel times to a financial center are more affected by investor sentiment than their counterparts, suggesting that the time cost of inter-city travel is an important consideration for financial market participants.

In addition, I try to seek additional evidence for the distance effect. First, I provide evidence that investor concentration matters for the distance effect. I find that among firms headquartered in the province, the three top-ranking firms in terms of stock market participation are more affected by investor sentiment than the other firms, which supports the notion that the concentration of individual investors matters for the investor sentiment effect.

Second, a huge strand of the literature documents that sell-side analysts show an optimistic bias in their earnings forecasts or stock recommendations and attribute it to their strategic behavior or cognitive bias (e.g., Brown, 1997; Lin and McNichols, 1998; Qian, 2009; Firth et al., 2013). Moreover, investor sentiment may affect analysts. Bagnoli et al. (2009) show that analysts pay attention to investor sentiment when issuing stock recommendations. Hribar and McNinnis (2012) find that analysts' earnings forecasts are

---

⁶“We believe HSR is not so competitive for distances over 1,500 kilometers” from Morgan Stanley Report “China High-Speed Rail” May 15, 2011.

more optimistic during periods with high market-wide sentiment. Corredor et al. (2014) confirm the cognitive bias in analyst forecasts with the finding that investor sentiment is positively and significantly related to analyst forecast errors in European markets. Consistent with this strand of literature, I find that analysts cover more local than nonlocal firms given the convenience of time and transportation costs. When investor sentiment is high, analysts issue more “Strong buy” or “Buy” recommendations. On average, local firms have 37 more “Strong buy” recommendations and 24 more “Buy” recommendations than nonlocal firms each quarter. Most importantly, analysts increase the frequency of favorable recommendations for local firms when investor sentiment is high. The results suggest that financial analysts tend to be more optimistic about local firms than nonlocal firms when investor sentiment is high, which may also drive the documented distance effect.

Furthermore, I provide evidence that there exists no difference between local and nonlocal firms in terms of institutional ownership, which helps to rule out the possibility that the geographic component in the sentiment-driven stock price relationship is induced by institutional investors and it further supports the notion that the geographically-distributed retail investors who are relatively inexperienced and more prone to market sentiment than institutional investors drive our main result.

Last but not least, I also offer evidence that the relation between geographic distance and the investor sentiment effect is not driven by strong economic performance and a good

market environment. Specifically, there is no significant difference in the investor sentiment effect between firms headquartered in provinces with high GDP and firms located in other provinces, even after excluding firms located right in the financial center.

These results address the endogeneity concern and support the notion that a geographic component exists in the sentiment-driven stock price relationship.

My studies contribute to the literature on the investor sentiment effect by adding an important geographic component. Unlike studies that focus on firm fundamental and corporate transparency in driving the sentiment effect, my study offers evidence that geographic distance from a firm's headquarters to its financial center can help to explain the sensitivity of stock prices to investor sentiment. Studies of the investor sentiment effect tend to construct a market-wide sentiment index for a stock market or a country. The geographic difference in the investor sentiment effect documented here raises the question of whether a regionally based sentiment index may be more relevant for investigating this effect in a large country such as China. This is an important issue for further research.

My study also contributes to the literature on the role of geographic distance in financial markets in two ways. First, studies have predominantly focused on mature markets such as those in the U.S. and other developed countries (e.g., Grinblatt and Keloharju, 2001; Butler, 2008; Almazan et al., 2010; John et al., 2011). Fewer studies have been conducted

on emerging markets where the development of financial markets (such as corporate governance, the information environment, investor protection, and education) and the investor base (with predominantly inexperienced retail investors), are vastly different from those of developed markets (Lemmon and Lins, 2003; Leuz et al., 2003). I focus on the world's largest emerging market and obtain evidence that is consistent with the salient characteristics of emerging markets. Second, my study also adds to the line of studies attempting to identify and distinguish the mechanisms underlying the geographic patterns of stock prices (e.g., Coval and Moskowitz, 2001; Pirinsky and Wang, 2006; Anand et al., 2011). By focusing on whether geographic proximity mitigates or intensifies investor sentiment effects, my study offers new evidence that the geographic pattern is driven by cognitive bias rather than information.

In addition to academic value, my study has policy implications and practical relevance for investors. Regarding the policy implications, for a large country such as China, the issue of whether to have just a few large financial centers or a network of decentralized financial centers has been intensively debated. My findings show that the existence of several large financial centers with a high concentration of financial institutions and investors may fuel the investor sentiment effect, especially in the context of an emerging market. They also show that investment in transportation infrastructure can help to reduce the effective distance and mitigate the investor sentiment effect.

Regarding practical relevance, my findings suggest that investors can make use of the

geographical component in investor sentiment effect to formulate their investment decisions. For example, investors can ride the sentiment effect by tilting their portfolios toward local stocks in the early stage of the high investor sentiment period and short-sell them in the later stage.

Chapter 2. Research background

Studies have documented that investor sentiment affects stock returns, although classical financial theory argues that rational arbitrageurs could bet against the risk of noise traders and correct for the mispricing they induce (e.g., Fama, 1965; Black, 1986; Brown and Cliff, 2005). A branch of the literature also offers evidence that the location of firms plays an important role in stock pricing (e.g., Anand et al., 2011; Garcia and Norli, 2012). While a firm's fundamentals and corporate transparency could explain the sensitivity of stock prices to investor sentiment (e.g., Baker and Wurgler, 2006; Firth et al., 2014), little attention has been paid to the location of the firm in the sentiment-driven stock price relationship. In this study, I investigate the influence of geographic distance on the sentiment-driven stock price relationship.

2.1. Investor sentiment and asset valuation

Investor sentiment, also known as “noise trader sentiment,” “the propensity of speculation” or “the optimism or pessimism about the overall stock market,” refers to the unpredictable opinions of noise traders about asset returns that are not warranted by fundamentals (De Long et al., 1990, hereafter DSSW; Baker and Wurgler, 2006). Classical financial theories argue that noise traders who make destabilizing speculations always lose money to rational arbitrageurs and will finally disappear from the financial market (Friedman 1953). However, DSSW (1990) refute that noise traders' beliefs may not revert to their mean

over a longer period and may become more extreme in the meantime, and hence that rational arbitrageurs lose money by trading against noise traders. As risk-averse arbitrageurs have reasonably short horizons, their willingness to arbitrage is limited. DSSW (1990) further argue that the unpredictability in noise traders' opinions of future stock prices deter arbitrage and affect prices even when there is no fundamental risk.

A large strand of the literature documents that investor sentiment can affect stock prices. Lee et al. (1991) suggest that the fluctuations in closed-end fund discounts are driven by changes in investor sentiment because the optimism or pessimism of individual investors affects the prices of underlying securities. Barberis et al. (1998) model investor sentiment and point out that conservatism and the representative heuristic make investors underreact and overreact to news, respectively. Using the number of "bullish," "bearish" or "neutral" newsletters to measure investor sentiment, Brown and Cliff (2005) provide evidence that investor sentiment drives stock prices to deviate from their fundamental values. When beginning-of-period investor sentiment is high, future stock returns are low as the stock price reverts to its fundamental value and vice versa. This investor sentiment effect changes according to the subjectivity of the valuation, limits on arbitrage, short-sale constraints and corporate information environment. Baker and Wurgler (2006, 2007) construct a market-level investor sentiment index and offer systematic evidence that investor sentiment can affect the cross-section of stock returns. Specifically, they suggest that stocks that are hard to value and difficult to arbitrage are more vulnerable to investor sentiment, and empirically document that small, young, high-volatility, unprofitable, non-

dividend-paying, extreme-growth and distressed stocks are affected more by investor sentiment. Stambaugh et al. (2012) suggest that the short-sale constraints contribute to overpricing when market-wide sentiment is high by exploring 11 asset pricing anomalies. Firth et al. (2015) demonstrate that firms with higher corporate transparency, measured by a bunch of variables including state ownership, the prevalence of related party transactions, accrual-based earnings management, audit opinions and the quality of audit firms are less affected by investor sentiment than firms with low corporate transparency. Cornell et al. (2014) further argue that high-quality accounting information reduces sentiment-related mispricing.

More recent studies seek new ways to measure investor sentiment. For example, Arif and Lee (2014) argue that aggregate corporate investment is an alternative measure of market-level investor sentiment. Da et al. (2015) suggest that households' daily Internet search volume can reveal market level investor sentiment. Aboody et al. (2016) offer empirical evidence that overnight stock returns have the same characteristics as the existing sentiment measure and propose that they could measure firm-specific investor sentiment. I follow the approach of Baker and Wurgler (2006) in constructing the market-level investor sentiment index for two reasons. First, the sentiment index produced by Baker and Wurgler is the most widely used to date (e.g., McLean and Zhao, 2014; Antoniou et al., 2015). Second, Firth et al. (2015) have already validated the index in the Chinese stock market by extending Baker and Wurgler's (2006) approach.

2.2. Firm location and asset pricing

The location of a firm matters for equity pricing. Coval and Moskowitz (2001) establish a geographic link between a fund investment and its performance. They find that fund managers earn substantial abnormal returns in their proximate investments and show that local investor ownership of a firm is positively related to its future expected returns after controlling for other factors that are known to capture a sizable portion of the cross-sectional variation in stock returns. Pirinsky and Wang (2006) document a strong co-movement in the stock returns of firms located in the same geographic area and argue that the price formation of the stock market has an important geographic component that is linked to the trading patterns of local residents. Hong et al. (2008) relate the stock price of a firm to the market conditions of its home locale by suggesting that the ratio of the aggregate book value of all firms headquartered in a certain region to the aggregate income of all households in the region has a negative effect on stock prices. Baik et al. (2010) find that both the level of and change in local institutional ownership predict future stock returns, and stocks in the local institutional investors' portfolio can earn higher excess returns around future earnings announcements than their counterparts. Anand et al. (2011) explore the importance of geographic proximity between firms and market makers for price discovery in the NASDAQ stock market and find that geographically proximate market makers contribute more to the price discovery of a firm's stock than distant market makers. Garcia and Norli (2012) show that firms with business activities concentrated in a small geographic area have higher returns than geographically dispersed

firms. El Ghouli et al. (2013) provide robust evidence that firms located in nonfinancial centers exhibit a higher cost of equity capital than firms headquartered in financial centers. Kubick and Lockhart (2016) find that a firm's proximity to a Securities Exchange Commission office influences the individual stock price crash risk. Overall, this branch of the literature highlights the importance of geography in asset pricing.

2.3. Geographic distance and information asymmetry

Another large strand of the literature demonstrates that geographic proximity helps to exploit local knowledge and gather information. As a result, being close can lower the level of information asymmetry between economic agents who possess a different information set. Coval and Moskowitz (2001) suggest that fund managers can earn abnormal returns on their geographically proximate investments with the help of their local information advantage. Ivković and Weisbenner (2005) argue that individual investors may find it easier to gather accurate value-relevant information on firms located close to them than on remote firms, and differentiate the “asymmetric information hypothesis” from the “familiarity hypothesis” by empirically documenting that proximate investments with value-relevant information can realize superior returns. Malloy (2005) finds that geographically proximate analysts are more accurate than other analysts because of their information advantage. Butler (2008) investigates municipal bond offerings and also finds that the comparative advantage of local investment banks is the most prominent for the most opaque issuers, which suggests that local underwriters have

easy access to specific information about issuers and thus can better evaluate them. Kang and Kim (2008) demonstrate that block acquirers who enjoy an information advantage (e.g., lower monitoring and time costs; more private and valuable information in terms of local targets) can perform better during the post-acquisition period. Baik et al. (2010) offer empirical evidence that the ownership of institutional investors can only predict future stock returns for local firms, and further highlight the informational advantages of local institutional investors. Bernile et al. (2015) suggest that geographical variation in firms' economic activities generates location-based information asymmetry among investors, which in turn influences their portfolio decisions and performance. Overall, these findings confirm that geographic distance may measure the level of information asymmetry by showing that economic agents can obtain more accurate information as distance erodes.

2.4. Geographic distance, familiarity and social interaction

The basic assumption behind the asymmetric information hypothesis⁷ is that the information disseminated near proximate regions is accurate and authentic. However, geographic proximity also exposes stocks to a wide variety of individual and institutional characteristics in certain regions at the same time. Specifically, physical proximity also facilitates social interactions and the transmission of sentiment and non-market information (Pirinsky and Wang, 2010). Another growing area of the literature suggests

---

⁷ The hypothesis that geographic distance could measure the level of information asymmetry.

that investors' preference for local investments is driven by a behavioral bias that makes them feel more comfortable investing in firms with which they are familiar. Based on a series of behavioral experiments, Heath and Tversky (1991) suggest that people prefer to bet on their own judgment in a context where they feel competent, and that this feeling is enhanced by familiarity and experience. They further argue that the reason investors with knowledge of diversification concentrate on a small number of firms is that investors regard themselves as capable of judging local investments. Using a Finnish dataset, Grinblatt and Keloharju (2001) empirically document that a firm's distance, language and culture are three important factors for investors' local preference because these factors enhance familiarity. Huberman (2001) also suggests that familiarity represents investors' belief that they currently have superior information or will obtain more information, and this familiarity drives their local investments. Seasholes and Zhu (2010) offer systematic empirical evidence that local investments do not generate abnormal returns and conclude that local preference is not driven by an information advantage. Korniotis and Kumar (2013) use a novel demographic-based proxy for smartness and find that "smart" investors have an information advantage while "dumb" investors reflect a psychological bias, such as overconfidence and sensation seeking, in their local investments.

In addition, geographic proximity may promote social interactions, and the actions of a reference group can affect an individual's preference (Horst and Scheinkman, 2006). Shiller (1984) suggests that investors' behavior is socially influenced and argue that fluctuations in attitude often occur in the population and often without any apparent

logical reason. Hong et al. (2004) find that households that interact with their neighbors or attend church are substantially more likely to participate in the stock market than other households. They also document that a fund manager is more likely to hold a specific stock if other managers in the same city invest in this firm and point out that investors spread stock market information by word of mouth (Hong et al., 2005). Brown et al. (2008) also suggest that an individual's investment decisions on stocks are affected by their neighbors via word-of-mouth. Baker et al. (2012) argue that social interaction is a mechanism used to spread sentiment. Han and Hirshleifer (2016) model how the bias toward sharing positive information induces more trading. These social interactions between financial market participants may increase the degree of familiarity and boost rumors, especially for unsophisticated individual investors who reveal cognitive bias.

2.5. Investor sentiment and sell-side analysts

Sell-side analysts play an important role in financial markets. They are in charge of synthesizing and disseminating information in capital markets, and investors with limited time or ability to analyze financial data often rely on analysts' work (Bradshaw, 2011). However, analysts show an optimistic bias in earnings forecasts and stock recommendations (e.g., Brown, 1997; Lin and McNichols, 1998; Qian, 2009; Firth et al., 2013), even though an analyst is more likely to turn over if his or her forecast accuracy is lower than that of his or her peers (Mikhail et al., 1999). Strategic behavior or cognitive bias contributes to this optimistic bias. For strategic behavior, Lim (2001) argues that

analysts trade off positive bias to improve management access and forecast accuracy because unfavorable forecasts limit or eliminate analysts' access to insiders. Hong and Kubik (2003) document that analysts who issue relatively optimistic forecasts are more likely to get favorable job separation, namely, moving upward in the hierarchy of brokerage firms because optimistic forecasts generate more trading and thus higher compensation for those analysts. In addition, Friesen and Weller (2006) model cognitive bias in analysts' earnings forecasts and find both US and international evidence that analysts are overconfident about the precision of their own information and suffer from cognitive dissonance bias. Mokoaleli et al. (2009) also suggest that the new buy stock recommendations reflect both behavioral bias and conflicts of interest in analysts.

In addition, analysts are affected by investor sentiment when they issue earnings forecasts and stock recommendations. Bagnoli et al. (2009) argue that analysts, who know that investor sentiment can affect stock prices, may consider both firm fundamentals and investor sentiment when making their recommendations. They provide empirical evidence that analysts pay attention to investor sentiment when issuing stock recommendations, and those that are positively correlated with investor sentiment tend to be less profitable. Hribar and McNinnis (2012) suggest that analysts' earnings forecasts are a direct measure of investors' earnings expectations and offer evidence that their forecasts are more optimistic during periods with high market-wide investor sentiment for uncertain and hard-to-value firms. Corredor et al. (2014) confirm that both cognitive bias and strategic behavior exist in analyst forecasts and find that investor sentiment is

positively and significantly related to analyst forecast errors in European markets.

2.6. Institutional background and geographic distribution in China

China became the world's second largest economy in terms of GDP in 2011 with persistent economic growth over the past 30 years. In addition, the Chinese stock market has grown exponentially to become the second largest in the world (after the United States).⁸ China has the third largest geographic area in the world, with three financial centers: Shanghai, Shenzhen and Beijing.

Shanghai's role as a predominant financial hub in the Asia Pacific region dates back to the 1930s. The first stock exchange market in China, the Shanghai Stock Exchange, was established in 1990 and has grown rapidly in the past decade. The Shanghai Stock Exchange provides a platform to raise capital for firms and offers investment opportunities for individuals. Shanghai also has a large urban economy and is a center for manufacturing and commerce.

In addition, the Lujiazui Financial District⁹ and Shanghai Free Trade Zone¹⁰ offer strong policy-supported advantages for financial market development in Shanghai, and

---

⁸ As of March 2017, the total market value of the Shanghai and Shenzhen exchanges was US\$7830.9 billion.

⁹ The state council declared the opening of Lujiazui district in 1990; financial firms in this district enjoy special tax policies, incentives and support, including personnel training and expedited visa services.

¹⁰ The state council launched the Shanghai Free Trade Zone in 2013. The zone cancels out a number of financial requirements for setting up a company in China and also introduces a simplified procedure for foreign investors to establish a company.

consequently, 445 multinational corporation (MNCs) had established regional headquarters there by 2013. Moreover, in the same year, Shanghai ranked first with respect to the total number of listed companies, listed H-shares companies, fund companies' headquarters, security companies' headquarters, total stock turnover value and so forth (Zhao et al., 2013).

Shenzhen is positioned as a regional financial center according to “The Outline of the Plan for the Reform and Development of the Pearl River Delta” announced by the National Development and Reform Commission. The geographic proximity to Hong Kong provides Shenzhen with a good foundation of specialized manufacturing and finance. The small and medium enterprise (SME) board and the ChiNext¹¹ on the Shenzhen Stock Exchange (SZSE) give small but high-growth and high-tech companies opportunities to raise capital. The ChiNext board has become the world's third largest market for growth enterprises and the largest among similar boards in Asia in terms of market capitalization and trade value (Zhao et al., 2013). Shenzhen also has a lot of investment and management companies, private equity and venture capital funds, which suits the needs of growing companies. Moreover, preferential government policies, including the establishment of Special Economic Zones and the Shenzhen-Hong Kong Modern Service Industry Cooperation Zone in Qianhai, have attracted foreign direct investment to start businesses and set up headquarters in Shenzhen.

---

¹¹ China's NASDAQ-style stock market.

Beijing is considered a financial center in China, although it does not have a stock exchange. A financial street was established by the State Council according to the 1993 Master Plan, and it was designed to host all regulatory agencies, including the central bank, the China Banking Regulatory Commission, the China Insurance Regulatory Commission and the China Securities Regulatory Commission, and headquarters of the big four state banks and non-banking enterprises (Chen and Chen, 2015). These financial regulatory organizations helped Beijing to become an information hinterland of financial policy and consequently attracted firms to become established there. Beijing's total A-share market capitalization ranked first in China and there are many large state-owned enterprises headquartered in Beijing. Beijing possesses the largest amount of A-share-raised capital and the second largest number of listed companies and trading value in China (Zhao et al., 2013). Moreover, Beijing is quite strong in the bond market, banking industry, PE fund market and insurance market and thus is an ideal place for MNC headquarters.

Overall, these three financial centers in China have a high concentration of financial institutions and a vast investor base, making them a potential origin of information and a venue for social interactions.

Chapter 3. Hypothesis Development and Empirical Results

3.1. Proximity and sentiment effect: more information or more behavioral bias?

Baker and Wurgler (2006) argue that firms that are more difficult to value are more affected by investor sentiment. In line with this view, Firth et al. (2015) find that firms with low corporate transparency are more vulnerable to market sentiment than their counterparts. Considering that the location of a firm influences equity pricing (e.g., Pirinsky and Wang, 2006; Garcia and Norli, 2012), it is plausible that a firm's proximity to a financial center might also affect its valuation difficulty and thus influence the sentiment-driven stock price relation. Money managers, major brokerage firms and thousands of investors are sited predominately near urban metropolitan areas (Loughran, 2007), especially in or near financial centers. Thus, a firm might be exposed to the general informational context or investors' characteristics in or near the financial center.

As discussed in the introduction, there are two competing hypotheses with regard to the impact of geographic distance on the investor sentiment effect. The first is the information hypothesis. Investors may obtain more value-relevant and firm-specific information on local firms than nonlocal firms. For a start, investors can easily access hard information on local firms. For instance, analysts who work for brokerage firms are more likely to cover local firms as the time and travel costs are low. O'Brien and Tan (2015) suggest that analysts are 80% more likely to cover IPO firms headquartered in their home states

than those in other states after controlling for industry specialization and underwriter status. Consequently, local firms are likely to have more frequent analyst forecasts, recommendations and reports than nonlocal ones. In addition, analysts who conduct corporate site visits can gather more firm-specific information and have greater forecast accuracy than other analysts (Cheng et al., 2016), which supports the notion that local firms have more hard information than do nonlocal firms. Furthermore, proximity offers an ideal channel for the collection of soft information. Agarwal and Hauswald (2010) find that borrower proximity facilitates the collection of soft information and the firm-bank distance drives a trade-off in the availability and pricing of credit. In a similar vein, John et al. (2011) argue that a firm's distance from shareholders limits the observability of managerial investment decisions, thereby exacerbating agency costs and strengthening the need for dividend payouts in the presence of free cash flow. If geographic proximity facilitates the diffusion of firm-specific information (including both hard and soft information), firms with greater proximity to a financial center should face a lower level of information asymmetry. With less valuation difficulty and more information to rely on, external investors have lower monitoring and information-gathering costs for local than nonlocal firms. As a result, investors depend less on their subjective judgment when investing in local firms, which are then less affected by investor sentiment than nonlocal firms.

The second hypothesis is based on the cognitive bias perspective. First, investors have a preference for firms they are familiar with. Huberman (2001) finds that clients of the

Regional Bell Operating Companies are much more likely to own stocks in the telephone company that provides their service than another telephone company, and argue that people prefer investing in the familiar. Grinblatt and Keloharju (2001) document that investors are more likely to hold, buy and sell the stocks of firms that are located close to them, communicate in their native tongue or have chief executives of the same cultural background. In addition, investors with limited cognitive ability may rely on familiarity as a screening device and therefore overweight familiar local stocks (Bernile et al., 2015). Huberman and Regev (2001) find that optimistic public attention induces a rise in the stock price of a biotechnology company even though no new information has been presented. Barber and Odean (2008) provide strong evidence that investors with scarce attention are net buyers of attention-grabbing stocks, including stocks that are in the news, are experiencing higher abnormal trading volume or have extreme one-day returns. I argue that investors who have limited attention may choose to invest in the familiar. Moreover, Seasholes and Zhu (2010) offer systematic empirical evidence that local portfolios do not generate abnormal returns, and purchases of local stocks significantly underperform sales of local stocks, suggesting that investors do not have value-relevant information about local stocks.

Financial centers have a high concentration of market participants, who are more familiar with nearby firms than those located far away. As a result, local firms are expected to be more affected by the cognitive bias of familiarity than nonlocal firms. Furthermore, I expect the familiarity effect to be stronger (weaker) during a period of high investor

sentiment. Because market participants are more (less) likely to talk about firms for investment when market sentiment is high (low) and such social interactions may affect the behavior of the market participants. For example, financial analysts may be more (less) willing to cover local firms when investor sentiment is high (low) and individual investors are more (less) likely to buy local firms when investor sentiment is high (low). As documented by Hong et al. (2004), social interactions can boost stock market participation. Consequently, local firms may be more affected by investor sentiment than nonlocal firms.

The influence of geographic proximity on the investor sentiment effect is therefore theoretically uncertain. Given that the information environment in China is poor and that various market participants (financial analysts and investors) are relatively inexperienced, I nevertheless formulate a hypothesis in favor of the cognitive bias perspective for empirical testing. The hypothesis is formally stated as follows.

All else being equal, local firms are more affected by investor sentiment than nonlocal firms.

3.2. Construction and validation of an investor sentiment index in the Chinese stock market

I construct an investor sentiment index to examine the sentiment effect in the Chinese

stock market. Baker and Wurgler (2006) develop an investor sentiment index using principal component analysis based on six underlying proxies for sentiment: closed-end fund discounts, market turnover, the number of IPOs, the average first-day returns on IPOs, the equity share of new issues and the dividend premium. Based on Baker and Wurgler's (2006) work, Firth et al. (2015) form a component sentiment index in the Chinese stock market by considering another two underlying proxies for investor sentiment: the deposit savings growth rate and the number of new investment accounts. For the deposit savings growth rate, Firth et al. (2015) point out that Chinese investors allocate their money mainly between banks and equity markets because of the underdeveloped nature of financial markets and capital control in China. When investors become disappointed with the stock market, they aggressively switch their investments in equities to their traditional concentration in savings accounts. Burdekin and Redfern (2009) find that the rising stock market sentiment exerted a significant negative effect on time deposit growth in China during 2003-2007, which supports the argument of Firth et al. (2015). For the number of new investments accounts, they suggest that China's stock market has a relatively short history compared with the developed financial markets, and that net inflows of new investors have entered the Chinese stock market in the last few decades. Due to their lack of investment experience and suffering during market downturns, investors may rely heavily on market sentiment rather than the rational calculation of firm fundamentals to participate in the stock market, especially in bull periods. Therefore, Firth et al. (2015) argue that the number of new investment accounts in the stock market may serve as a potential proxy for investor sentiment. In this paper, I

follow the approach used by both Baker and Wurlger (2006) and Firth et al. (2015) and construct a monthly investor sentiment index in the Chinese stock market from January 2003 to December 2015.

3.2.1. Construction of the investor sentiment index in the Chinese stock market

Following Firth et al. (2015), I construct an investor sentiment index based on the variations of seven underlying proxies: the closed-end fund discount (CEFD), market turnover (TURN), the number of IPOs (NIPO) and their average first-day return (RIPO), the proportion of equity issues in the total issue of equity and long-term debt (Eshare), the deposit savings growth rate (DSG) and the number of new investment accounts (NACT). The first five variables are constructed in the same way as described by Baker and Wurlger (2006). Specifically, CEFD refers to the value-weighted average difference between the net asset value (NAV) per share and market price of each closed-end fund share, divided by the net asset value (NAV) per share. TURN is constructed as the natural log of the ratio of the amount of monthly market turnover trading to the aggregate market value, detrended by the 5-month moving average. NIPO is the number of IPOs and RIPO refers to the market-adjusted first-day returns. Eshare is the proportion of new equity issuance to total capital (total equity issuances plus bank borrowing) raised in a year. DSG is defined as the residual from the regression of the growth rate of seasonally adjusted deposit savings on the growth rate of M0. NACT is defined as the natural log of the number of new investment accounts.

Following Baker and Wurlger (2006), I isolate the common component in the aforementioned seven proxies using principal component analysis, considering that each proxy is likely to include a sentiment-based and an idiosyncratic or non-sentiment-related component. As some proxies may take longer to reveal the same sentiment, I perform a factor analysis on all of the variables and their lags to determine the best lead-lag structure for each proxy. First, I estimate the first principal component of the seven proxies and their lags. This procedure produces a first-stage index with 14 loadings, one for each of the current and lagged proxies. I then calculate the correlation between the first-stage index and the current and lagged values of each proxy. Finally, I construct the sentiment index as the first principal component based on the correlation matrix of the seven proxies; I use the current or lagged value of the seven proxies, depending on which has a higher correlation with the first-stage index. I standardize the second-stage sentiment index so that the index has unit variance. The first principal component explains 37.37% of the sample variance, which is comparable with the explanatory power of the sentiment index reported in major developed countries (Baker et al., 2012). The correlation of the first- and second-stage indexes is 0.96, suggesting that little information is lost in dropping the seven proxies with other time scripts. The sentiment index is constructed as follows:

$$\begin{aligned} sentiment_t = & -0.1207CEFD_t + 0.5317TURN_{t-1} + 0.3607NIPO_t + \\ & 0.1895RIPO_{t-1} + 0.4217Eshare_t - 0.2844DSG_{t-1} + 0.5272NACT_{t-1} \end{aligned} \quad (1)$$

As the variations of some proxies may have both a sentiment component and a business cycle component, I then construct a second index that explicitly removes the business

cycle variation from each of the proxies before the principal component analysis. Following Firth et al. (2015), I regress each of the seven raw proxies on the industrial production and consumption growth rates. The residuals labeled with a superscript $\perp$ from these regressions are cleaner proxies for investor sentiment, and I construct the sentiment index following the same procedures as before. The corresponding first principal component explains 36.87% of the sample variance, which is still comparable with the explanatory powers of the sentiment index reported in major developed countries (Baker et al., 2012). The final investor sentiment index is constructed as follows:

$$\begin{aligned} sentiment_t^\perp = & -0.0702CEFD_t^\perp + 0.5174TURN_{t-1}^\perp + 0.3134NIPO_t^\perp + \\ & 0.2115RIPO_{t-1}^\perp + 0.4482Eshare_t^\perp - 0.3352DSG_{t-1}^\perp + 0.5207NACT_{t-1}^\perp \end{aligned} \quad (2)$$

Similar to the results of Firth et al. (2015), NACT has the highest loadings in the sentiment index and the factor loading on DSG is higher than those on CEFD, NIPO and RIPO, which indicates that both proxies are important for constructing an investor sentiment index in the Chinese stock market. Table 1 presents the summary statistics and correlations between the components of the sentiment index. All of the components are significantly correlated in the expected directions. When investor sentiment is high, the market trading volume moves up, the number of IPOs and the average first-day return of IPOs increase, the proportion of new equity issuances to total capital goes up and the number of new investment accounts soars, whereas the money-supply-adjusted deposit savings and the closed-end fund discount decrease. NACT and TURN are significantly correlated with every other component, consistent with the notion that NACT and TURN

make the largest contribution to the investor sentiment index. The correlations between the orthogonal variables only change slightly.

3.2.2 Validation of the investor sentiment index in Chinese stock market

To validate the investor sentiment index, I apply the same approach as Firth et al. (2015). Specifically, the general sentiment effect implies that if excessively optimistic sentiment drives the stock price up from its fundamental value, periods of high sentiment should be followed by low future stock returns as the stock price reverts back to its fundamental value. To investigate the sentiment effect, I regress the future excess market returns ($ExRET_{t+k}$) on the sentiment index SENT and hypothesize that the coefficients of the sentiment index should be significantly negative. I compute $ExRET_{t+k}$ as the value-weighted market returns minus the risk-free rate¹² and cumulate it over various horizons from 1 to 24 months. Following the asset pricing literature (e.g., Brown and Cliff, 2005; Lemmon and Portniaguina, 2006), I use macroeconomic variables as controls, namely, the contemporaneous market return ($ExRET_t$), aggregate earnings-to-price ratio (EP_t), monthly interest rate (RF_t) and inflation rate ($INFL_t$). However, the investor sentiment index I construct is predetermined¹³ and not strictly exogenous,¹⁴ and the estimation on the lagged endogenous variables from the ordinary least squares regression is biased in finite samples (Stambaugh, 1999). In addition, the residuals from the regressions with

---

¹² Measured by the monthly interest rate on one-year fixed deposits.

¹³ The past market sentiment affects current investor behavior but has little impact on future investor sentiment. For instance, a past financial crisis will lower the intention to invest in stock market now, but future financial growth or a hot market in the near future will not affect market-wide investor sentiment.

¹⁴ The seven proxies for investor sentiment have an impact on stock returns.

overlapping observations can be serially correlated up to lag K-1 under both the null and alternative hypotheses, which fully account for time-varying expected returns (Swaminathan, 1996). To cope with these problems, I follow the approach of Nelson and Kim (1993) and Neal and Wheatley (1998) in testing significance by comparing the test statistics with their empirical distributions computed from randomization simulations.

The regression model is as follows:

$$\sum_{k=1}^K [ExRET_{t+k}/K] = \alpha + \beta SENT_t + \gamma_1 ExRET_t + \gamma_2 EP_t + \gamma_3 RF_t + \gamma_4 INFL_t + \varepsilon_{t+k} \quad (3)$$

where SENT refers to the investor sentiment index and *ExRET* refers to the equal-weighted market return over the risk-free rate. The summary statistics of SENT and other control variables are shown in Panel C of Table 1. Panel A of Table 2 presents the coefficients and simulation p-values from the time-series regressions of Equation (3). The coefficients on SENT decline monotonically up to the 12th month with a little rebound thereafter, and become significantly negative at a p-value of 0.05 or lower as the forecasting horizon extends to 11 months and beyond. Panel B of Table 2 illustrates the trend of the SENT coefficients over time. These results are similar to those of Firth et al. (2015)¹⁵ and validate the sentiment index as an important determinant of stock price. Furthermore, Figure 1 compares my investor sentiment index with that of Firth et al. (2015). The two sentiment index lines are almost overlapping from January 2003 to December 2009, providing direct evidence of the validity of the investor sentiment index

---

¹⁵ They document that the coefficients on SENT decline monotonically until the fourteenth month and become significantly negative at a p-value of 0.05 or lower as the forecasting horizon extends to 12 months and beyond.

in my study.

3.3. Empirical analysis

3.3.1. Sample description and data source

Table 3 shows the sample selection procedures. The initial sample contains all of the firms that have been listed for at least three years on either the Shanghai or the Shenzhen A-share exchanges (i.e., firms listed before the end of 2013). I exclude firms listed on the Growth Enterprise Market because they are relatively small. I also exclude financial firms and delisted firms. The final sample includes 2,058 firms listed on either the Shanghai or Shenzhen exchange market. I hand collect the addresses of firms' headquarters and obtain longitude and latitude data at the prefecture level from Google map. My investigation period covers January 2003 to December 2015 and the investor sentiment index has 156 monthly observations. All financial data and stock price data come from the China Stock Market and Accounting Research (CSMAR) database.

3.3.2. Distance and the investor sentiment effect

I begin by examining the difference in investor sentiment effect between local and nonlocal firms. And I also exclude firms located exactly in financial centers and retest the main hypothesis to differentiate the effect of financial centers from the impact of

geographic distance in the sentiment-driven stock price relationship. Two mechanisms are related to this issue. First, if the empirical result is consistent with the information hypothesis, one may expect the relation between geographic distance and the investor sentiment effect to hold even when firms located in financial centers are excluded from our sample because information, especially soft information, may not be confined to the financial center. Second, if the result is in line with the behavioral bias perspective, the interaction between distance and the investor sentiment effect will not disappear. Social interactions and the degree of familiarity may not be confined to the financial center, and there may be some spillover of investor bias in nearby areas. In any case, the relation between proximity and investor sentiment still holds after excluding firms located in financial centers.

I measure the geographic distance between a firm's headquarter and the center of whichever financial center (Beijing, Shanghai or Shenzhen) is closest. Specifically, firms that are headquartered within a certain radius (250 miles, 100 miles and 100 kilometers) of Beijing, Shanghai or Shenzhen are considered as local firms whereas firms that are headquartered outside the radius are considered nonlocal firms. I also compare the investor sentiment effect on firms located within a radius of 100 kilometers and firms outside the radius of 1,000 kilometers to offer clearer evidence of the difference in investor sentiment effect between local and nonlocal firms. While the investor sentiment effect becomes statistically significant from the 11-month horizon onwards, I focus on the time horizons of 9, 12 and 15 months to observe the association between sentiment

and subsequent returns. Following Brown and Cliff (2005), I investigate the sensitivity of stock prices to investor sentiment after controlling for the earnings-to-price ratio (EP), risk free rate (RF), inflation rate (INFL) and the three Fama-French factors (MKT, SMB and HML). The regression model is as follows:

$$\sum_{k=1}^K ExRET_{t+k} = \alpha + \beta SENT_t + \gamma_1 EP_t + \gamma_2 RF_t + \gamma_3 INFL_t + \gamma_4 MKT_t + \gamma_5 SMB_t + \gamma_6 HML_t + \varepsilon_{t+k} \quad (4)$$

where $\sum_{k=1}^K ExRET_{t+k}$ is the K-period portfolio return; SENT is the investor sentiment index; EP refers to the market-level earnings-to-price ratio; RF is the risk-free rate measured by the monthly interest rate on one-year fixed deposits; INFL refers to the monthly inflation rate; MKT is the excess return on market portfolio; SMB refers to the average return on small stock portfolios over the average returns on large stock portfolios; and HML refers to the average return on high book-to-market stock portfolios over the average returns on high book-to-market stock portfolios. I expect β to be significantly negative as it represents the sensitivity of subsequent returns to investor sentiment. I apply the same approach as in Equation (3) to cope with the biases in the test statistics. To test the main hypothesis, I form equal-weighted portfolios on local and nonlocal firms. To distinguish the effect of geographic distance from that of financial centers, I form equal-weighted portfolios on local and nonlocal firms after excluding firms located in Beijing, Shanghai or Shenzhen.

Table 4 summarizes the number of firms in each portfolio using different geographic

cutting boundaries with or without firms located in Beijing, Shanghai or Shenzhen, respectively. I carry out the regression model estimation for the portfolios sorted by geographic distance. Table 5 reports the SENT coefficient estimation results on local and nonlocal firms, with the local portfolio consisting of firms headquartered within a radius of 250 miles, 100 miles or 100 kilometers¹⁶ from Beijing, Shanghai or Shenzhen, whereas the nonlocal portfolio consists of firms that are headquartered outside these boundaries. D1 consists of the firms located within a radius of 100 kilometers from Beijing, Shanghai or Shenzhen, whereas D3 consists of firms located outside a radius of 1,000 kilometers from Beijing, Shanghai or Shenzhen. For the sake of brevity, I do not report the coefficients for EP, RF, INFL, MKT, SMB and HML because they are not the primary focus of my study.¹⁷ Consistent with the previous results of the price reversal shown in Panel B of Table 2 and the results reported by Firth et al. (2015) on corporate transparency, the bias-adjusted coefficient estimates of SENT over the three horizons are all negative and significant over 12- and 15-month horizon. Most importantly, in line with the behavioral bias hypothesis, local firms are more affected by investor sentiment than nonlocal firms (the distance effect). The differences between the SENT coefficients in local and nonlocal firms are statistically significant across all three horizons for all of the aforementioned geographic cutting boundaries. The results are also significant in economic terms. For instance, an increase of one standard deviation in the sentiment index results in an additional decline of 150 basis points in the stock prices of local firms over

---

¹⁶ I also test the difference between local and nonlocal firms in terms of sentiment effect using 150miles, 200miles as local and nonlocal firms cutting boundary, the results are similar.

¹⁷ The detailed regression results are available from the author upon request.

nonlocal firms in the subsequent 12 month-period using the geographic cutting boundary of 250 miles.

Table 6 summarizes the SENT coefficient estimation results for local and nonlocal firms after excluding firms located in Beijing, Shanghai or Shenzhen. Firstly, although the bias-adjusted coefficients on SENT are still negative and significant for all firms over 12 and 15 months, the differences between the SENT coefficients in local and nonlocal firms are marginally significant when using 250 miles and 100 miles as the geographic cutting boundary. The difference in sentiment coefficients between local and nonlocal firms is almost zero and insignificant when using 100 kilometers as cutting boundary, but the number of nonlocal companies are over ten times more than local companies. Secondly, the magnitude of the differences also declines. For example, local firms bear an additional decrease of 100 basis points¹⁸ in their market value compared with nonlocal firms for a one standard deviation increase in market optimism for the 12-month horizon when using 100 miles as local and nonlocal cutting boundary. In addition, the differences in the SENT coefficients between the D1 and D3 portfolios are marginally significant for the above three horizons, and the magnitude of the difference in sentiment effect between D1 and D3 portfolio also decreased. For instance, D1 firms bear additional 130 basis points decrease of their market value than D3 firms in the following 12 months for one standard deviation increase in market sentiment without firms headquartered in Beijing, Shanghai or Shenzhen, as compared to a decrease of 290 basis points when including firms in

---

¹⁸ The additional decrease is 170 basis points when including firms located in Beijing, Shanghai or Shenzhen, see Panel B of Table 4.

Beijing, Shanghai or Shenzhen.

I interpret these results as indicating that financial centers play an important but not indispensable role in the relation between geographic distance and the investor sentiment effect. Specifically, a financial center offers facilities and intermediaries that facilitate social interactions in the financial market and therefore increase the level of familiarity with local firms. However, investors' behavioral bias may not be confined within the financial center but may spill out to nearby areas. As a result, the distance effect remains after excluding firms located within financial centers.

3.3.3. Chinese High Speed Railway reform as an exogenous shock

Next, I use the Chinese HSR reform as an exogenous shock to investigate whether the construction of transport infrastructure affects the impact of geographic distance on the sentiment-driven stock price relationship. HSR construction is intended to connect provincial capitals and other major cities by faster means of transportation, and the decision on the placement of HSR lines is built on a comprehensive consideration of the economic development, population and resource distribution, national security, environmental concerns and social stability of each region (Mid-to-Long Railway Plan, 2008), which have no direct relation with the investor sentiment effect. Therefore, the HSR reform serves as a reasonable exogenous shock in this setting. Moreover, the upgrading of transportation reduces both the time and transportation cost of inter-city

travel. As a result, it may facilitate social interactions between market participants and thus increase the degree of familiarity. If this were the case, I would expect the distance effect to decrease or even disappear for firms with headquarters in cities connected to HSR lines, whereas it should still exist for firms unconnected to HSR lines.

First, I examine the difference in the sentiment effect on nonlocal firms unconnected to and connected to HSR lines, respectively, before and after the launch of HSR services, respectively. I pay special attention to whether nonlocal firms are connected to HSR lines because the magnitude of the decrease in time and transportation costs is lower for local firms. In order to mitigate the confounding effect of aviation, I restrict nonlocal firms located within 1500 kilometer-radius of Beijing, Shanghai or Shenzhen and do the same analysis as before. Second, I investigate whether the distance effect exists if I divide firms into local and nonlocal groups based on travel times by railway, airline and toll road from the headquarters of a firm to a financial center.

Specifically, I separate 2,058 firms into local and nonlocal portfolios using 250 miles as the geographic cutting boundary. This procedure yields 1,205 local firms and 853 nonlocal firms. Considering that local firms are within the 250-mile radius of either Beijing, Shanghai or Shenzhen and the lowest speed of HSR lines is 250 km/h, the HSR reform has much less impact on shortening the travel time for local firms than for nonlocal firms, I then focus on the effect of HSR connection for nonlocal firms. Specifically, I

choose 2007¹⁹ as the event year because China's HSR service opened on April 18, 2007. I hand collect the data for the specific year in which a prefecture-level city opened its HSR service from the China Railway Press and match it with the headquarters of nonlocal firms. This procedure yields 215 nonlocal firms connected to the HSR in 2007 and 230 nonlocal firms that remain unconnected in 2007 and afterwards. Following Chang and Wong (2009), I match firms that experienced the HSR shock in 2007 with firms that did not. I first match each shocked firm with a firm in the same industry that stayed unconnected to the HSR in 2007 and afterwards. If multiple firms satisfy this criteria, then I include the firm closest in size to that of the shocked firm in year -1. I find 205 unconnected firms that are matched with shocked firms on the basis of industry and firm size.

Applying the same approach taken for Equation (4), I sort the portfolios on local and nonlocal firms, respectively. While the group of local firms remains unchanged, nonlocal firms are divided into two groups according to the foregoing procedure. Specifically, I compute the difference in equal-weighted portfolio returns between local and nonlocal firms unconnected to HSR lines (control group) and between local and nonlocal firms that experienced HSR reform (treatment group), respectively. Then, I investigate the difference in the relation between distance and the investor sentiment effect on the treatment and control groups both before and after 2007. In addition, I restrict nonlocal

---

¹⁹ I also try to use different years as the event year. However, there are only 31 nonlocal firms and 50 nonlocal firms that began connected to HSR in 2008 and 2009, respectively. The results are less meaningful due to small sample size.

firms located within 1500 kilometer-radius of Beijing, Shanghai or Shenzhen and do the same analysis on treatment and control group, respectively to mitigate the confounding effect of aviation.

Moreover, I divide firms into local and nonlocal groups according to the travel times from a firm's headquarters to a financial center. The measurement of travel times is based on a comprehensive consideration of railway, airline and toll road. Firstly, for firms located within 300 kilometer-radius of Beijing, Shanghai or Shenzhen, I assume that market participants can reach financial centers by toll road within 4 hours²⁰ since the average speed of the National Trunk Highway System is 100 km/h (Li and Shum, 2001) and considering that there may exist some curves on toll road because the construction of toll road is based on the consideration of regional development²¹; Secondly, I hand collect the shortest travel time from the headquarter of a firm to Beijing, Shanghai or Shenzhen via railway (including HSR) from Chinese Railway Press from 2006 to 2013. I add one hour to all the travel time by train for market participants to travel to railway station and come to their destination. Finally, for firms located outside 1500 kilometer-radius²² of Beijing, Shanghai or Shenzhen, I hand collect the travel time from a firm to the nearest financial center by plane from Ctrip. I also add 3 hours for all the travel time by plane since passengers are supposed to arrive in airport 2 hours before the airplane take off and I leave 1 hour for them to come to airport and get to their destination. Firms located in the

---

²⁰ A one-day round business trip.

²¹ I leave 100 kilometers for the curve.

²² The competitive radius of HSR is within 1500 kilometer-radius.

cities within 12 hours to Beijing, Shanghai or Shenzhen are grouped as local firms whereas firms headquartered in the cities exceed 12 hours to Beijing, Shanghai or Shenzhen are classified as nonlocal firms. In order to better test the role of travel time in sentiment-driven stock price relationship, I further group firms located in the cities within 4 hours to Beijing, Shanghai or Shenzhen as local firms whereas firms headquartered in the cities exceed 12 hours to Beijing, Shanghai or Shenzhen are classified as nonlocal firms. I then examine whether the distance effect exists for local and nonlocal firms based on travel time following the same approach as for Equation (4).

Table 7 reports the SENT coefficient estimation results for the treatment group. Panel A shows that the distance effect exists before the HSR reform. Specifically, local firms are more affected by investor sentiment than nonlocal firms from 2003 to 2006 across three time horizons. For instance, an increase of one standard deviation in the sentiment index results in a decline of an additional 390 basis points in the stock prices of local over nonlocal firms in the subsequent 12-month period. However, the distance effect disappears after the HSR reform in 2007. Panel B of Table 7 shows that the differences in the SENT coefficient estimation between local and nonlocal firms becomes insignificant, although the biased-adjusted SENT coefficients are universally and significantly negative for the three horizons. Moreover, the distance effect for the treatment group also disappears for the whole sample period from 2003 to 2015, as reported in Panel C of Table 7.

In contrast, in the control group, local firms are more affected by investor sentiment than nonlocal firms across the three sample periods, namely, 2003 to 2006, 2007 to 2015 and 2003 to 2015, as reported in Table 8, Panels A to C. For example, local firms bear a decrease of an extra 290 basis points in their market prices compared with nonlocal firms for each one standard deviation increase in investor optimism for the 12-month horizon before the HSR reform in 2007, and this distance effect is also significant from 2007 to 2015. Local firms bear an additional 320-point decline in their market value over nonlocal firms in the subsequent 15 months for each one standard deviation increase in the sentiment index over the whole sample period from 2003 to 2015. Obviously, the distance effect holds for the control group when nonlocal firms remain unconnected to the HSR.

Furthermore, Figure 2 presents the trends of the differences in the SENT coefficients over a 24-month period on the time-series regressions of Equation (4) before and after the HSR reform for the treatment group and control group, respectively. Panel A shows that the trends for the two groups are quite similar. Both of the lines are negative over the 24-month period, suggesting that local firms are more affected by the general market sentiment than nonlocal firms in both the treatment and control groups before the HSR reform. However, after the reform, the gap between the SENT coefficients of the two groups becomes larger. Specifically, the differences in the SENT coefficients for the treatment group are generally positive, indicating that nonlocal firms are more affected by investor sentiment, although this relation is insignificant. In contrast, the differences in the SENT coefficients for the control group are generally negative, suggesting that the

distance effect still exists after the HSR reform for control group.

Table 9 and Table 10 conduct the same analysis as in Table 7 and Table 8 but restrict the firms headquartered within 1500 kilometer-radius of Beijing, Shanghai or Shenzhen. The results are robust in the competitive radii of HSR.

The above results provide support for the notion that the construction of transportation infrastructure may alter the effect of geographic distance on the sentiment-driven stock price relationship. To put it another way, I verify the importance of geographic distance in explaining the sensitivity of stock prices to market sentiment using the HSR reform as an exogenous shock to test the changes in the distance effect on the treatment and control groups.

Table 11 reports the results for the relation between travel time and sentiment effect. Firms with shorter travel times to a financial center are more affected by investor sentiment than their counterparts. Panel A shows that firms within a 12-hour travel distance of a financial center are more affected by investor sentiment than firms with a travel distance of more than 12 hours to either Beijing, Shanghai or Shenzhen. Panel B summarizes that firms within a 4-hour travel time of a financial center are more affected by investor sentiment than firms with a travel time of more than 12 hours to either Beijing, Shanghai or Shenzhen. For both Panel A and Panel B, the SENT coefficients become more negative as the travel time from a firm's headquarters to the nearest financial center

decreases. For instance, the SENT coefficients for firms within a 4-hour travel time are considerably lower than those of their control counterparts and their differences are all statistically significant at the 5% level across the three horizons. These results provide important evidence that travel time is a channel to stimulate social interactions across cities and firms located in cities near financial center bear higher degree of familiarity than other firms.

3.3.4. Additional analyses

3.3.4.1. Stock market participation and investor sentiment effect

To offer more evidence for the behavioral bias hypothesis, I examine the relation between the stock market participation rate and investor sentiment effect. Brown et al. (2008) establish a casual relation between an individual's decision to own a stock and the average stock market participation rate of the individual's community. They further argue that word-of-mouth communication drives this causal effect. This study supports the notion that social interactions affect the regional stock market participation rate. The mechanism by which the market participation rate influences the sentiment-driven stock price relation is as follows. Provinces with a high stock market participation rate are likely to have more investors talking about their investments, which could increase the degree of familiarity. For instance, investors may envy others' success in the stock market and follow their neighbors or acquaintances in buying certain stocks when they have little information

about the fundamentals of these stocks. Therefore, firms headquartered in provinces with a high market participation rate are affected more by investor sentiment than firms headquartered in low market participation regions. I conduct the following test to support my argument.

I divide firms into portfolios based on the stock market participation rate of the province where a firm's headquarters is located from 2003 to 2014.²³ Specifically, the stock market participation rate is defined as the provincial-level total stock accounts divided by the year-end population in that province. Firms located in the top three²⁴ stock market participation provinces are identified as high stock market participation firms and the rest as low stock market participation firms each year. I apply the same approach taken for Equation (4) to compare the difference in the investor sentiment effect for High firms and Low firms.

Panel A of Table 12 lists the top three provinces in terms of the stock market participation rate from 2003 to 2014. The table shows that Shanghai always ranks first, ranging from 43.79% in 2005 to 80.90% in 2009, followed by Beijing, Guangdong and Tianjin. In addition, the participation rate in these provinces increases across years, suggesting the development of the financial market in China. Panel B of Table 12 reports the regression results for the relation between the investor sentiment effect and the stock market

---

²³ Before April 2015, a person could only open only one stock trading account in China. In most cases, investor's account address and living address are same.

²⁴ I consider the top three provinces because the stock market participation rate in other provinces is quite low, around 10% or lower from 2003 to 2014.

participation rate. Across three horizons, I consistently find that the SENT coefficients on high stock market participation firms are more negative than those on their counterparts. The High-minus-Low portfolio comparison suggests that the differences in the SENT coefficients are highly significant. This result supports the behavioral bias hypothesis by offering evidence that firms located in the provinces with high market participation are more affected by investor sentiment than firms located in other provinces.

3.3.4.2. Number of analyst recommendations and the investor sentiment effect

In addition, I seek additional evidence to explain why firms that are headquartered proximate to a financial center are more affected by market sentiment than firms that are farther away. I pay special attention to sell-side analysts because they are in charge of offering investment advice to their clients and generating more trades for brokerage firms, and because their workplaces are predominantly located in or near financial centers. Thus, analysts may conveniently choose to cover more local than nonlocal firms given the time and transportation costs. Moreover, previous studies document that sell-side analysts reveal an overoptimistic bias in their earnings forecasts or stock recommendations and attribute it to their strategic behavior or cognitive bias (e.g., Brown, 1997; Lin and McNichols, 1998; Qian, 2009; Firth et al., 2013). Most importantly, investor sentiment can affect analysts. Bagnoli et al. (2009) suggest that analysts who know that investor sentiment has an impact on stock prices pay attention to market sentiment when issuing stock recommendations. If this is the case, analysts' behavioral bias may fuel the degree

of familiarity and thereby intensify the sentiment effect near financial centers.

Following Hribar and McNnis (2012), I calculate the analysts' forecast errors on local and nonlocal portfolios, respectively, and compare the difference in analysts' optimistic bias between local and nonlocal firms. However, the difference is not statistically significant.²⁵ A plausible explanation is that analysts know that unsophisticated individual investors pay more attention to the frequency of recommendations rather than the earnings forecasts and use investors' psychological bias to fuel the investor sentiment effect on local firms.

I then examine the impact of the number of analyst recommendations on the sentiment-driven stock price relationship. Following Bergman and Roychowdhury (2008), I investigate the association between sentiment and number of analyst recommendations on local and nonlocal firms using the following regression model:

$$FREQ_{it} = \alpha + \beta_1 Local_i + \beta_2 SENT_{t-1} + \beta_3 Local_i * SENT_{t-1} + \gamma X_{t-1} + \delta Z_{it-1} + \varepsilon_{it} \quad (5)$$

The dependent variable $FREQ_{it}$ is the quarterly stock recommendation frequency issued for firm i in quarter t . I compute the quarterly recommendation frequency based on three types, namely, "Strong buy", "Buy" and "Hold or worse," and run the preceding regression for each of these three types. $SENT_{t-1}$ is the average monthly investor

---

²⁵ The result is available upon request.

sentiment index in quarter $t-1$ and Local is a dummy variable that equals 1 if the firm is local and 0 otherwise. X_{t-1} represents a vector of quarterly control variables including the seasonally adjusted percentage change in GDP in quarter $t-1$ (GDPCHG), gross returns on the value-weighted market index over the 12 months before the beginning of quarter t (PASTMARKET), and the volatility of gross market returns during the 12 months before the beginning of quarter t (MARKETVOL). Z_{it-1} stands for a vector of firm-quarter control variables that contains the natural logarithm of a firm's asset value Size, book-to-market ratio (BTM) and gross market returns over the past 12 months (RET), and the standard deviation of each firm's monthly returns over the past 12 months (RETVOL). I also include lagged capital expenditure as a percentage of total assets (CAPEX) and lagged return on assets (ROA) as control variables. All of the firm-level control variables are winsorized at the 1% level to reduce the effect of outliers. I include the binary indicator variables for firms that report a loss in any of the previous four quarters (LOSS) and for firm quarters containing equity issues (ISSUE). I further include the institutional ownership (IOHLD) of each firm to control for the institutional attention. Finally, I include firm and year fixed effects because the number of analyst recommendations increases during my sample period, as shown in Panel B of Table 13. All of the analyst recommendation and stock price data are taken from CSMAR and the GDP data are from the Statistical Bureau of the PRC from 2003 to 2015.

Panel A of Table 13 reports the univariate test results for quarterly analyst coverage on local and nonlocal firms from 2003 to 2015. Analyst coverage equals 0 if no analyst is

following a firm in a certain quarter. Analyst coverage for local firms is then computed as the number of analysts who issued stock recommendations for firms located within a 250-mile radius of Beijing, Shanghai or Shenzhen, and that for nonlocal firms as the number of analysts who issued stock recommendations for firms headquartered outside this radius. The univariate analysis shows that more analysts follow local than nonlocal firms, and the difference is statistically significant at the 1% level. Although the magnitude is small, I attribute it to many of the 2,058 firms in my sample period receiving no analyst attention.

Table 14 summarizes the regression results for the relation between the number of stock recommendations and investor sentiment for local and nonlocal firms. The summary statistics of the variables in the regression are presented in Panel A of Table 14. The mean quarterly frequency of “Strong buy” recommendations in the sample is 3.16, and the average quarterly frequencies of “Buy” recommendations and “Hold or worse” recommendations are 4.45 and 0.89, respectively. The average number of stock recommendations in each quarter is 8.5. While the average firm is followed by two analysts each quarter, the mean institutional ownership of the firms in the sample is 2.28%. On average, 57% of firms are classified in the local group in my sample. Panel B of Table 14 reports the initial results for the relationship between stock recommendations and investor sentiment in terms of local and nonlocal firms. On average, local firms have 41 more “Strong buy” recommendations than nonlocal firms in each quarter. This number is significant in economic terms because the mean number of “Strong buy”

recommendations in the sample is 3.16. The local firms also have 27 more “Buy” recommendations than nonlocal firms on average. Moreover, the significantly positive interaction of Local and SENT for “Strong buy” recommendations indicates that such recommendations are more likely for local firms when investor sentiment is high.

However, analysts self-select whether to follow a firm. Rather than local firms’ geographic location, these results could be driven solely by local firms being followed by more overoptimistic analysts, making the firms subject to more overoptimistic recommendations. To address this problem, I use Heckman’s two-stage method to address the concern of a self-selection bias. First, I use the following probit regression to estimate the probability of a firm receiving a stock recommendation:

$$Treated_{it} = \alpha + \beta_0 NUM_{it} + \beta_1 Local_i + \beta_2 SENT_{t-1} + \beta_3 Local_i * SENT_{t-1} + \gamma X_{t-1} + \delta Z_{it-1} + \varepsilon_{it} \quad (6)$$

where $Treated_{it}$ is a dummy variable that equals 1 if firm i has analyst recommendations in quarter t , and 0 otherwise. NUM_{it} is the number of firms in the same industry as firm i in quarter t . NUM_{it} is associated with the probability of issuing stock recommendations because the information gathering cost reduces as the number of firms in a certain industry increases, but it has no obvious relation to whether a firm receives a certain type of stock recommendation. This variable is used for identification purposes. The definitions of the other control variables are the same as those given for Equation (5). In the second stage, I use the following regression to re-estimate the relation

between stock recommendations and investor sentiment on local and nonlocal firms after controlling for macro- and firm-level variables by including the inverse Mills ratio obtained from the first-stage regression:

$$FREQ_{it} = \alpha + \beta_1 Local_i + \beta_2 SENT_{t-1} + \beta_3 Local_i * SENT_{t-1} + \gamma X_{t-1} + \delta Z_{it-1} + \beta_4 Lambda + \varepsilon_{it} \quad (7)$$

where *Lambda* is the inverse Mills ratio obtained from the first-stage regression. Panel C of Table 14 reports the result for Equation (6). The coefficient on *NUM_{it}* is significantly positive, suggesting that a firm is more likely to receive analyst coverage when there are more firms operating in the same industry. In addition, a firm is more likely to be covered by analysts when market sentiment, book-to-market ratio, return on assets, capital expenditure percentage, institutional ownership, past returns and past market volatility are high, and if the firm is large or located close to a financial center, while past return volatility, gross returns on market index and seasonally adjusted GDP are low or suffered a loss in the last quarter.

Panel D of Table 14 reports the regression results for Equation (7). Compared with the baseline model in Panel C of Table 14, where I did not correct for the self-selection bias, the main results still hold. Specifically, the coefficients of *SENT* for “Strong buy” and “Buy” recommendations are significantly positive, suggesting that analysts issue more positive stock recommendations when investor sentiment is high. Similar to the results for the baseline regression, local firms obtain more “Strong buy” recommendations and

“Buy” recommendations. When investor sentiment is high, local firms get more “Strong buy” recommendations.

In sum, these results suggest that local firms receive more favorable recommendations than nonlocal firms. When investor sentiment is high (low), local firms have more (fewer) “Strong buy” recommendations than nonlocal firms. The results from the Heckman two-stage model suggest that the overall results are unlikely to be driven by a self-selection bias.

Furthermore, I investigate the implications of stock recommendations on the value of local and nonlocal firms. Following Firth et al. (2013), I examine the stock market’s reaction to analyst recommendations upon report issuance. Specifically, I focus on the cumulative abnormal returns (CARs) centered on the recommendation date over the three-day event window from one day before until one day after. I use the market-adjusted abnormal returns, defined as the stock’s return over the value-weighted market returns for firms listed on the Shanghai and Shenzhen stock exchanges. I then cumulate the daily abnormal returns over the three-day event window to get the three-day CARs ($CAR(-1,+1)$). Panel E of Table 14 shows that positive recommendations, including “Strong buy” and “Buy” recommendations, generate significantly positive three-day CARs for both local and nonlocal firms, whereas negative “Hold or worse” recommendations produce significantly negative three-day CARs for both local and nonlocal firms, suggesting that favorable recommendations have a positive effect on short-term stock prices while

unfavorable ones have a negative effect on short-term prices. In addition, the abnormal returns for the three types of recommendations on local firms are all significantly greater than those for nonlocal firms, suggesting that market reactions to the stock recommendations for local firms are more optimistic than those for nonlocal firms, which is consistent with the conjecture that the familiarity bias is initiated near financial centers and financial analysts fuel investors' behavioral bias.

3.3.4.3 Institutional ownership and investor sentiment

Previous literatures document that local institutional investors can utilize informational advantages and realize abnormal returns in their local investments (e.g.: Coval and Moskowitz, 2001; Baik et al., 2010). As institutional investors are mainly located near financial centers, some may argue that local firms may have higher institutional ownership than nonlocal firms and institutional investor may play a role in the relation between distance and sentiment-driven stock price relationship.

I do a two-way and a four-way analysis to address this concern. Table 15 reports a comparison of institutional holdings between local and nonlocal firms in high- and low-sentiment period. Specifically, Panel A of Table 15 shows that there is no significantly differences between local and nonlocal firms in terms of institutional ownership. In Panel B, when investor sentiment is high, institutional investors buy more stocks indiscriminate of local and nonlocal firms than low-sentiment period. For example, institutional

ownership of high-sentiment period is 0.75%²⁶ more than that of low-sentiment period for nonlocal firms. However, the difference of institutional holding between local and nonlocal firms within each sentimental group is not statistically significant. These results help to relief the concern that the geographic component in the sentiment-driven stock price relationship is driven by institutional investors and supports the notion that it is the retail investors who are relatively inexperienced and more prone to market sentiment may drive our main result.

3.3.4.4. GDP and the investor sentiment effect

Some may argue that the distance effect could be caused by different levels of economic development rather than cognitive bias as I propose, as regions located near to financial centers also tend to be more developed economically. For example, investors who reside in more economically developed regions tend to be richer and may be more willing to take risks or even to gamble when investing in the stock market. If this is the case, one may expect firms headquartered in more developed areas to be more affected by investor sentiment than other firms.

To address this concern, I investigate the relation between GDP and the investor sentiment effect. I divide firms into portfolios based on the nominal GDP of the province where a firm's headquarters is located from 2003 to 2015. Firms located in the top 10

---

²⁶ This magnitude is significant since the mean of institutional holding is 2.28%.

GDP provinces are identified as High GDP firms and those in other provinces as Low GDP firms in each year. I apply the same approach as for Equation (4) to compare the difference in the investor sentiment effect on High and Low portfolios.

Panel A of Table 16 presents the regression results for the relation between GDP and the investor sentiment effect. I find the consistent result that the SENT coefficients on high GDP firms are no more negative than those on low GDP firms across three horizons. The High-minus-Low portfolio comparison suggests that the differences in SENT coefficients are not significant. Panel B of Table 16 conducts the regression analysis on GDP and the sentiment effect after excluding firms located in either Beijing, Shanghai or Guangdong (where Shenzhen is located) and the results are similar across three horizons. These results support the notion that my main result on the distance effect is not driven by different levels of economic development.

Chapter 4. Conclusions

This study examines the role of geographic distance in the sentiment-driven stock price relationship in the context of China's stock market. I provide comprehensive evidence that firms with greater proximity to financial centers are more affected by the general market sentiment than firms far away from financial centers. In addition, using the Chinese HSR reform as an exogenous shock, I find that the relation between distance and the investor sentiment effect disappears for firms that become connected to HSR lines but

holds for firms that stay unconnected, suggesting that upgrading the transport infrastructure facilitates social interactions and mitigates the importance of geographic distance in the sentiment-driven stock price relation. Moreover, I find that firms within a 4-hour travel distance to a financial center are more affected by sentiment than those that are more than 12 hours away. The result confirms that travel time is an important determinant of the investor sentiment effect. Finally, I offer additional evidence for the relation between geographic distance and the sensitivity of stock price to investor sentiment. First, firms located in provinces with higher stock market participation rates are more affected by investor sentiment than firms headquartered in other provinces. Second, analysts increase the frequency of favorable recommendations for firms that are located closer to financial centers when investor sentiment is high and the stock market reactions to analyst recommendations are in line with my findings. Finally, I provide evidence that the distance effect is not driven by institutional investors or regional economic development. Overall, my study highlights the importance of geographic distance in explaining investor sentiment effects.

Table 1: Summary statistics of investor sentiment data and control variables

This table reports the summary statistics for the sentiment component from 2003 to 2015. CEFD is the value-weighted average discount rate of closed-end mutual funds. TURN is the natural log of monthly market trading volume divided by the average market value, detrended by the 5-month moving average. NIPO is the monthly number of initial public offerings. RIPO is the average market-adjusted first-day returns of initial public offerings. Eshare is the ratio of equity issues to total equity issues and bank borrowing. DSG is the growth rate of savings deposits controlling for the growth of money supply (M0). NACT is the natural log of new investment accounts (in thousands). SENT is the first principal component of the seven sentiment proxies.

Panel A reports the summary statistics for the raw sentiment component. In Panel B, I regress each of the seven proxies on the growth in industrial production and consumption growth rate. The orthogonalized proxies, prefixed with “residual,” are the residuals from these regressions. SENT_1 is the first principal component of the seven orthogonalized proxies. In Panel C, EP is the market-level earnings-to-price ratio. RF is the risk-free rate, measured by the monthly interest rate on one-year fixed deposits. INFL is the monthly rate of inflation. MKT is the excess return on the market portfolio of stocks. SMB (HML) is the difference each month between average returns on the small-stock portfolio (the high BM portfolio) and average returns on the large-stock portfolio (the low BM portfolio). *** denotes significance at the 0.01 level, ** at the 0.05 level and * at the 0.1 level.

Panel A: Raw data

Variable	Mean	SD	P25	Median	P75
CEFD_t (%)	17.85	12.28	6.74	16.43	26.03
TURN_t-1	6.10	0.48	5.75	6.03	6.42
NIPO_t	7.64	7.25	0.00	6.00	13.00
RIPO_t-1(%)	0.77	0.77	0.36	0.58	0.92
Eshare_t(%)	1.34	1.45	0.42	0.94	1.78
DSG_t-1(%)	0.05	0.88	-0.39	0.06	0.46
NACT_t-1	6.22	1.27	5.11	6.26	7.04

Correlation	CEFD	TURN	NIPO	RIPO	Eshare	DSG	NACT	SENT	SENT_1
CEFD_t (%)	1								
TURN_t-1	-0.0627	1							
NIPO_t	-0.2629***	0.2460***	1						
RIPO_t-1(%)	0.2417***	0.2124**	-0.1740*	1					
Eshare_t(%)	0.0331	0.4302***	0.4492***	0.4015***	1				
DSG_t-1(%)	0.0964	-0.1680**	-0.1740**	-0.1173	-0.2740***	1			
NACT_t-1	-0.4975***	0.7467***	0.4434***	0.1518	0.4052***	-0.2030**	1		
SENT	-0.1952**	0.8600***	0.5834***	0.3065***	0.6822***	-0.4601***	0.8527***	1	
SENT_1	-0.1018	0.7683***	0.4944***	0.3235***	0.6872***	-0.5378***	0.6997***	0.9102***	1

Panel B: Controlling for macroeconomic conditions

		Variable	Mean	SD	P25	P50	P75		
		residual_CEFD_t (%)	0.00	10.09	-6.27	-3.73	6.94		
		residual_TURN_t-1	0.00	0.45	-0.30	-0.09	0.32		
		residual_NIPO_t	0.00	7.00	-6.10	-1.08	4.25		
		residual_RIPO_t-1(%)	0.00	0.74	-0.38	-0.09	0.11		
		residual_Eshare_t(%)	0.00	1.39	-0.78	-0.34	0.47		
		residual_DSG_t-1(%)	0.00	0.88	-0.43	0.03	0.40		
		residual_NACT_t-1	0.00	1.07	-0.63	-0.12	0.55		
Correlation	R_CEFD	R_TURN	R_NIPO	R_RIPO	R_Eshare	R_DSG	R_NACT	SENT	SENT_1
R_CEFD_t (%)	1								
R_TURN_t-1	0.0793	1							
R_NIPO_t	-0.3861***	0.2155***	1						
R_RIPO_t-1(%)	0.1616*	0.2111**	-0.2095**	1					
R_Eshare_t(%)	-0.0739	0.4455***	0.4073***	0.3549***	1				
R_DSG_t-1(%)	0.1002	-0.1877**	-0.1956**	-0.1397	-0.3036***	1			
R_NACT_t-1	-0.4462***	0.7005***	0.4356***	0.1834*	0.4389***	-0.2591***	1		
SENT	-0.1291	0.7809***	0.4846***	0.2857***	0.6334***	-0.4787***	0.7538***	1	
SENT_1	-0.1128	0.8313***	0.5035***	0.3398***	0.7201***	-0.5386***	0.8365***	0.9102***	1

Panel C: SENT and control variables

Variable	Mean	SD	P25	P50	P75
SENT	0.000	1.000	-0.677	-0.073	0.592
EP (%)	0.089	0.031	0.062	0.083	0.106
RF (%)	0.224	0.054	0.188	0.210	0.250
INFL (%)	2.761	2.112	1.400	2.400	4.000
MKT	0.009	0.089	-0.051	0.015	0.055
SMB	0.009	0.047	-0.015	0.011	0.036
HML	0.001	0.033	-0.017	0.000	0.019

Correlation	SENT	EP	RF	INFL	MKT	SMB	HML
SENT	1						
EP (%)	-0.1224	1					
RF (%)	-0.0239	-0.2915***	1				
INFL (%)	-0.0645	-0.371***	0.6688***	1			
MKT	0.1929**	0.2192***	-0.1881**	-0.2712***	1		
SMB	0.1211	-0.1646**	0.0091	-0.0126	0.1715**	1	
HML	-0.0694	0.0976	0.0333	0.0304	0.0903	-0.4872***	1

Table 2: Validation of the sentiment index and its relation with future stock returns

This table reports the results from the time-series regression of excess market returns on investor sentiment over a 24-month period in Equation (3). Panel A presents the coefficients from the time-series regressions; the simulated p-values reported in parentheses are for two-sided tests of the hypothesis that the slope coefficient does not equal 0. Panel B shows the graph of the SENT coefficient β over time.

Panel A: Time-series regressions of excess market returns on investor sentiment (SENT)

K	1	2	3	4	5	6	7	8
$\beta \cdot 100$	0.95	0.84	0.33	-0.01	-0.20	-0.23	-0.43	-0.64
p-value	(0.19)	(0.19)	(0.18)	(0.82)	(0.95)	(0.72)	(0.56)	(0.33)
K	9	10	11	12	13	14	15	16
$\beta \cdot 100$	-0.84	-0.99	-1.09	-1.12	-1.09	-1.12	-1.10	-1.05
p-value	(0.20)	(0.10)	(0.05)	(0.03)	(0.01)	(0.01)	(0.00)	(0.00)
K	17	18	19	20	21	22	23	24
$\beta \cdot 100$	-0.99	-0.96	-1.01	-1.00	-0.98	-0.98	-0.97	-0.96
p-value	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

Panel B: Graph of the SENT coefficient β over time

Figure 1: Comparison of sentiment indices

This graph compares the investor sentiment indices. sentiment1 (dotted line) is the sentiment index constructed by Firth et al. (2015), whereas sentiment2 (solid line) represents the sentiment index formulated in this study.

Table 3: Sample selection

This table presents the sample selection procedure for firms listed on China's A-share market.

	No. Observations
1. All firms listed on the A-share market before the end of 2013	2,549
2. Dropping firms listed on the Growth Enterprise Market	2,194
3. Dropping financial firms	2,148
4. Dropping delisted firms	2,058

Table 4: Number of firms in each geographic portfolio

This table summarizes the number of firms sorted in each portfolio using different geographic cutting boundaries. Panel A (Panel B) reports the number of firms in local and nonlocal portfolios including (excluding) firms located in Beijing, Shanghai or Shenzhen when applying 100kilometers, 100 miles and 250 miles as local and nonlocal cutting boundary; Panel C (Panel D) report the number of firms in D1, D2 and D3 portfolio including (excluding) firms headquartered in Beijing, Shanghai or Shenzhen, where D1 represents the number of firms located within a 100-kilometer radius of Beijing, Shanghai or Shenzhen, while D3 shows the number of firms headquartered outside a 1,000-kilometer radius of Beijing, Shanghai or Shenzhen and D2 stands for firms located between 100-kilometer radius of Beijing, Shanghai or Shenzhen and 1,000-kilometer radius of Beijing, Shanghai or Shenzhen.

Panel A: The number of firms on local and nonlocal portfolio including firms located in Beijing, Shanghai or Shenzhen

Cutting Boundary	Local firms	Nonlocal firms	Total
100kilometers	610	1448	2058
100miles	831	1227	2058
250miles	1205	853	2058

Panel B: The number of firms on local and nonlocal portfolio excluding firms located in Beijing, Shanghai or Shenzhen

Cutting Boundary	Local firms	Nonlocal firms	Total
100kilometers	108	1448	1556
100miles	329	1227	1556
250miles	703	853	1556

Panel C: The number of firms on D1, D2 and D3 portfolio including firms located in Beijing, Shanghai or Shenzhen

D1	D2	D3	Total
610	1213	235	2058

Panel D: The number of firms on D1, D2 and D3 portfolio excluding firms located in Beijing, Shanghai or Shenzhen

D1	D2	D3	Total
108	1213	235	1556

Table 5: Geographic distance and the investor sentiment effect

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are defined as for Equation 4. I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns on local firms and Nonlocal represents those on nonlocal firms. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Panel A uses 250 miles as the cutting boundary for local and nonlocal firms, Panel B uses 100 miles and Panel C uses 100 kilometers. D1 represents the average returns of firms located within a 100-kilometer radius of Beijing, Shanghai or Shenzhen, while D3 shows the average returns of firms headquartered outside a 1,000-kilometer radius of Beijing, Shanghai or Shenzhen and D2 stands for the average returns of firms located between 100-kilometer radius of Beijing, Shanghai or Shenzhen and 1,000-kilometer radius of Beijing, Shanghai or Shenzhen. Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms with 250 miles as the local–nonlocal boundary

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.065 (0.18)	-0.053 (0.32)	-0.012 (0.01)
12 months	-0.104 (0.04)	-0.090 (0.10)	-0.015 (0.03)
15 months	-0.116 (0.02)	-0.099 (0.07)	-0.017 (0.06)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms with 100 miles as the local–nonlocal boundary

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.069 (0.15)	-0.055 (0.30)	-0.014 (0.00)
12 months	-0.109 (0.03)	-0.091 (0.09)	-0.017 (0.01)
15 months	-0.120 (0.01)	-0.101 (0.06)	-0.019 (0.02)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms with 100 kilometers as the local–nonlocal boundary

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.071 (0.14)	-0.055 (0.28)	-0.015 (0.00)
12 months	-0.110 (0.03)	-0.093 (0.09)	-0.017 (0.02)
15 months	-0.122 (0.01)	-0.101 (0.05)	-0.019 (0.05)

Panel D: Equal-weighted portfolio regression on local and nonlocal firms with 100 kilometers and 1,000 kilometers as the boundary dividing D1, D2 and D3

Horizon	D1	D2	D3	D1-D3
9 months	-0.070 (0.13)	-0.058 (0.25)	-0.047 (0.39)	-0.024 (0.00)
12 months	-0.110 (0.02)	-0.096 (0.07)	-0.081 (0.13)	-0.029 (0.00)
15 months	-0.123 (0.01)	-0.105 (0.03)	-0.088 (0.08)	-0.034 (0.00)

Table6: Geographic distance and the investor sentiment effect (excluding firms located in financial center)

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4) after excluding firms located in Beijing, Shanghai or Shenzhen. I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns of local firms and Nonlocal those of nonlocal firms. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Panel A uses 250 miles as the cutting boundary for local and nonlocal firms, Panel B uses 100 miles and Panel C 100 kilometers; D1 represents the average returns of firms located within a 100-kilometer radius of Beijing, Shanghai or Shenzhen while D3 shows the average returns of firms headquartered outside a 1,000-kilometer radius of Beijing, Shanghai or Shenzhen, D2 stands for the average returns of firms located between 100-kilometer radius of Beijing, Shanghai or Shenzhen and 1,000-kilometer radius of Beijing, Shanghai or Shenzhen.. Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms with 250 miles as the local–nonlocal boundary

Horizon	Local	Nonlocal	Local- Nonlocal
9 months	-0.059 (0.23)	-0.053 (0.32)	-0.006 (0.11)
12 months	-0.098 (0.06)	-0.090 (0.10)	-0.008 (0.10)
15 months	-0.108 (0.03)	-0.099 (0.07)	-0.010 (0.13)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms with 100 miles as the local and nonlocal boundary

Horizon	Local	Nonlocal	Local- Nonlocal
9 months	-0.062 (0.20)	-0.054 (0.30)	-0.007 (0.01)
12 months	-0.101 (0.05)	-0.091 (0.09)	-0.010 (0.01)
15 months	-0.111 (0.03)	-0.099 (0.06)	-0.011 (0.02)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms with 100 kilometers as the local–nonlocal boundary

Horizon	Local	Nonlocal	Local- Nonlocal
9 months	-0.056 (0.26)	-0.056 (0.27)	0.000 (0.93)
12 months	-0.094 (0.07)	-0.093 (0.08)	-0.001 (0.92)
15 months	-0.104 (0.04)	-0.102 (0.05)	-0.001 (0.94)

Panel D: Equal-weighted portfolio regression on local and nonlocal firms with 100 kilometers and 1,000 kilometers as the boundary dividing D1, D2 and D3

Horizon	D1	D2	D3	D1-D3
9 months	-0.056 (0.26)	-0.058 (0.25)	-0.046 (0.41)	-0.009 (0.08)
12 months	-0.094 (0.06)	-0.097 (0.07)	-0.081 (0.13)	-0.013 (0.10)
15 months	-0.105 (0.03)	-0.105 (0.03)	-0.088 (0.08)	-0.016 (0.09)

Table 7: Geographic distance and investor sentiment effect for local and nonlocal firms connected to HSR lines

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4) for the treatment group (local and nonlocal firms that are connected to HSR lines). I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns on firms located within a 250-mile radius of Beijing, Shanghai or Shenzhen and Nonlocal represents the returns on firms outside these radii. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Panel A reports the equal-weighted portfolio regression on the treatment group before the HSR reform (2003-2006). Panel B reports the equal-weighted portfolio regression on treatment group after the HSR reform (2007-2015). Panel C reports the equal-weighted portfolio regression on the treatment group for the whole sample period (2003-2015). Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2003 to 2006

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	0.149 (0.01)	0.194 (0.01)	-0.045 (0.00)
12 months	0.104 (0.08)	0.144 (0.04)	-0.039 (0.00)
15 months	0.016 (0.94)	0.036 (0.75)	-0.019 (0.03)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2007 to 2015

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	-0.110 (0.01)	-0.115 (0.01)	0.005 (0.42)
12 months	-0.141 (0.00)	-0.143 (0.00)	0.003 (0.72)
15 months	-0.119 (0.00)	-0.120 (0.01)	0.001 (0.92)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2003 to 2015

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	-0.066 (0.18)	-0.059 (0.27)	-0.006 (0.51)
12 months	-0.105 (0.04)	-0.097 (0.09)	-0.007 (0.54)
15 months	-0.116 (0.02)	-0.108 (0.06)	-0.007 (0.65)

Table 8: Geographic distance and investor sentiment effect for local and nonlocal firms unconnected to HSR lines

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4) for the control group (local and nonlocal firms that are unconnected to HSR lines). I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns on firms located within a 250-mile radius of Beijing, Shanghai or Shenzhen while Nonlocal stands for the average stock returns on firms headquartered outside these radii. Local–Nonlocal represents the differences in the average stock returns between local and nonlocal firms. Panel A reports the equal-weighted portfolio regression on the control group before the HSR reform (2003-2006). Panel B reports the equal-weighted portfolio regression on the control group after the HSR reform (2007-2015). Panel C reports the equal-weighted portfolio regression on the control group for the whole sample period (2003-2015). Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2003 to 2006

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	0.149 (0.01)	0.177 (0.01)	-0.028 (0.00)
12 months	0.104 (0.08)	0.133 (0.04)	-0.029 (0.00)
15 months	0.016 (0.94)	0.041 (0.65)	-0.024 (0.00)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2007 to 2015

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	-0.110 (0.01)	-0.096 (0.02)	-0.014 (0.03)
12 months	-0.141 (0.00)	-0.124 (0.00)	-0.017 (0.08)
15 months	-0.119 (0.00)	-0.098 (0.02)	-0.022 (0.08)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2003 to 2015

Horizon	Local	Nonlocal	Local– Nonlocal
9 months	-0.065 (0.18)	-0.041 (0.49)	-0.024 (0.00)
12 months	-0.104 (0.04)	-0.076 (0.18)	-0.029 (0.00)
15 months	-0.116 (0.02)	-0.083 (0.13)	-0.032 (0.01)

Table 9: Geographic distance and investor sentiment effect for local and nonlocal firms connected to HSR lines for subsamples

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4) for the treatment group (local and nonlocal firms that are connected to HSR lines). I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns on firms located within a 250-mile radius of Beijing, Shanghai or Shenzhen and Nonlocal represents the returns on firms outside these radii but within 1500 kilometer-radius of Beijing, Shanghai or Shenzhen. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Panel A reports the equal-weighted portfolio regression on the treatment group before the HSR reform (2003-2006). Panel B reports the equal-weighted portfolio regression on treatment group after the HSR reform (2007-2015). Panel C reports the equal-weighted portfolio regression on the treatment group for the whole sample period (2003-2015). Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2003 to 2006

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	0.149 (0.01)	0.193 (0.01)	-0.045 (0.00)
12 months	0.105 (0.09)	0.144 (0.05)	-0.040 (0.00)
15 months	0.017 (0.94)	0.035 (0.74)	-0.019 (0.04)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2007 to 2015

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.111 (0.01)	-0.115 (0.01)	0.005 (0.41)
12 months	-0.140 (0.00)	-0.144 (0.00)	0.003 (0.73)
15 months	-0.119 (0.00)	-0.120 (0.01)	0.001 (0.91)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms connected to HSR lines from 2003 to 2015

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.065 (0.17)	-0.059 (0.27)	-0.006 (0.50)
12 months	-0.104 (0.04)	-0.096 (0.09)	-0.007 (0.55)
15 months	-0.115 (0.02)	-0.108 (0.06)	-0.007 (0.64)

Table 10: Geographic distance and investor sentiment effect for local and nonlocal firms unconnected to HSR lines for subsamples

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4) for the treatment group (local and nonlocal firms that are connected to HSR lines). I report the coefficient on SENT for each equal-weighted portfolio formed on geographic distance. Local represents the average stock returns on firms located within a 250-mile radius of Beijing, Shanghai or Shenzhen and Nonlocal represents the returns on firms outside these radii but within 1500 kilometer-radius of Beijing, Shanghai or Shenzhen. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Panel A reports the equal-weighted portfolio regression on the treatment group before the HSR reform (2003-2006). Panel B reports the equal-weighted portfolio regression on treatment group after the HSR reform (2007-2015). Panel C reports the equal-weighted portfolio regression on the treatment group for the whole sample period (2003-2015). Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2003 to 2006

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	0.149 (0.01)	0.189 (0.00)	-0.040 (0.00)
12 months	0.105 (0.09)	0.144 (0.03)	-0.041 (0.00)
15 months	0.017 (0.93)	0.049 (0.57)	-0.032 (0.00)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2007 to 2015

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.111 (0.01)	-0.100 (0.02)	-0.011 (0.08)
12 months	-0.141 (0.00)	-0.128 (0.00)	-0.013 (0.12)
15 months	-0.119 (0.00)	-0.100 (0.01)	-0.019 (0.09)

Panel C: Equal-weighted portfolio regression on local and nonlocal firms unconnected to HSR lines from 2003 to 2015

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.065 (0.18)	-0.043 (0.48)	-0.023 (0.00)
12 months	-0.104 (0.04)	-0.079 (0.18)	-0.026 (0.01)
15 months	-0.116 (0.02)	-0.086 (0.13)	-0.029 (0.03)

Figure 2: Sentiment coefficients over time before and after HSR reform

The figure shows the difference in the sentiment coefficients from Equation (4) between local and nonlocal firms over the 24-month period for the treatment group and the control group before and after the HSR reform. $Dif\_beta\_Unconnected$ (dotted line) is the difference in the sentiment coefficients between local and nonlocal firms that are not connected to an HSR line over the 24-month period and $Dif\_beta\_Connected$ is the difference between those that are connected to HSR lines over the 24-month period. Panel A shows the difference in sentiment coefficients between the treatment and control groups from 2003 to 2006; Panel B shows the difference in sentiment coefficients between the treatment and control groups from 2007 to 2015.

Panel A: Sentiment coefficients before the HSR reform, 2003-2006

Panel B: Sentiment coefficients after the HSR reform, 2007-2015

Table 11: Travel time and the investor sentiment effect

This table summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4). I report the coefficient on SENT for each equal-weighted portfolio formed on the travel times by train, plane and toll road. For Panel A, Local represents the average stock returns on firms headquartered in cities within a 12-hour travel time of Beijing, Shanghai or Shenzhen while Nonlocal represents the returns on firms headquartered in cities more than 12-hours away from Beijing, Shanghai or Shenzhen. For Panel B, Local represents the average stock returns on firms headquartered in cities within a 4-hour travel time of Beijing, Shanghai or Shenzhen while Nonlocal represents the returns on firms headquartered in cities more than 12-hours away from Beijing, Shanghai or Shenzhen. Local–Nonlocal represents the difference in the average stock returns between local and nonlocal firms. Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on local and nonlocal firms with 12 hours as local-nonlocal boundary

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.155 (0.00)	-0.140 (0.01)	-0.014 (0.00)
12 months	-0.242 (0.00)	-0.226 (0.00)	-0.016 (0.00)
15 months	-0.260 (0.00)	-0.241 (0.00)	-0.019 (0.00)

Panel B: Equal-weighted portfolio regression on local and nonlocal firms with 4 hours as local group and 12 hours as nonlocal group

Horizon	Local	Nonlocal	Local-Nonlocal
9 months	-0.156 (0.00)	-0.140 (0.01)	-0.015 (0.00)
12 months	-0.242 (0.00)	-0.226 (0.00)	-0.016 (0.01)
15 months	-0.260 (0.00)	-0.241 (0.00)	-0.019 (0.05)

Table 12: Stock market participation and the investor sentiment effect

This table presents the results of the effect of the stock market participation rate on the investor sentiment effect. Panel A summarizes the top three provinces in terms of stock market participation rates from 2003 to 2014. Panel B summarizes the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML), which are shown in Equation (4). I report the coefficient on SENT for each equal-weighted portfolio formed on the stock market participation rate. High represents the average stock returns on firms headquartered in provinces with a top three stock market participation rate while Low represents the returns on firms headquartered in other provinces. High–Low represents the differences in the average stock returns between high and low stock market participation firms. Simulated p-values are reported in parentheses and are for two-sided tests of the hypothesis that the slope coefficient β does not equal 0.

Panel A: Top three stock market participation provinces

2003		2004		2005		2006	
Province	MPR	Province	MPR	Province	MPR	Province	MPR
Shanghai	45.01%	Shanghai	46.83%	Shanghai	43.79%	Shanghai	44.31%
Beijing	23.08%	Beijing	24.15%	Beijing	24.69%	Beijing	20.87%
Tianjin	15.12%	Tianjin	15.20%	Tianjin	15.09%	Tianjin	14.01%

2007		2008		2009		2010	
Province	MPR	Province	MPR	Province	MPR	Province	MPR
Shanghai	68.44%	Shanghai	73.53%	Shanghai	80.90%	Shanghai	74.06%
Beijing	34.86%	Beijing	38.01%	Beijing	45.01%	Beijing	46.77%
Guangdong	18.82%	Guangdong	20.66%	Guangdong	23.43%	Guangdong	23.85%

2011		2012		2013		2014	
Province	MPR	Province	MPR	Province	MPR	Province	MPR
Shanghai	75.08%	Shanghai	75.04%	Shanghai	74.95%	Shanghai	78.03%
Beijing	48.22%	Beijing	48.22%	Beijing	48.33%	Beijing	49.62%
Guangdong	25.40%	Guangdong	25.93%	Guangdong	26.46%	Guangdong	27.57%

Panel B: Equal-weighted portfolio regression on high market participate portfolios and low market participate portfolios

Horizon	High	Low	High-Low
9 months	-0.073	-0.054	-0.019
	(0.13)	(0.3)	(0.00)
12 months	-0.115	-0.091	-0.023
	(0.03)	(0.09)	(0.00)
15 months	-0.127	-0.101	-0.027
	(0.01)	(0.06)	(0.00)

Table 13: Summary statistics on stock recommendations

This table presents the summary statistics on the number of analysts' stock recommendations from 2003 to 2015. Panel A reports the univariate analysis in terms of the quarterly number of recommendations on local and nonlocal firms (250 miles as the geographical cutting boundary). Panel B summarizes the yearly number of stock recommendations.

Panel A: The number of analysts for local and nonlocal firms

Variables	Obs	Mean
Number of analysts for local firms	46,238	2.128
Number of analysts for nonlocal firms	35,500	1.905
Local-nonlocal		0.223***
t-statistic		8.61

Panel B: The number of stock recommendations by year

Year	Strongbuy	Buy	Hold or worse	All
2003	303	641	1011	1955
2004	523	1546	1006	3075
2005	2062	5710	4291	12063
2006	4664	8795	3160	16619
2007	7108	10804	2343	20255
2008	15921	31751	11694	59366
2009	15773	38297	12900	66970
2010	30448	46610	8688	85746
2011	40009	55712	7940	103661
2012	33131	48906	7688	89725
2013	33810	46133	5710	85653
2014	36882	40347	3932	81161
2015	37521	28671	2251	68443
Total	258155	363923	72614	694692

Table 14: Number of analyst recommendations and the investor sentiment effect

This table presents the full results on the relation between the number of analyst recommendations and investor sentiment. Panel A presents the summary statistics for the variables in the analysis. “Strong buy” is the number of quarterly “Strong buy” recommendations, “Buy” is the number of quarterly “Buy” recommendations, “Hold or worse” is the sum of quarterly “Netural”, “Hold” and “Sell” recommendations, and “All” is the total stock recommendations issued in the corresponding quarter. SENT is the average monthly investor sentiment index in a quarter. Local is a dummy variable that equals 1 if the firm is identified as a local firm and 0 otherwise. GDPCHG is the seasonally adjusted percentage change in GDP in a quarter, PASTMARKET is the gross returns on the value-weighted market index over the 12 months prior to the beginning of a quarter and MARKETVOL is the volatility of gross market returns during the 12 months prior to the beginning of a quarter. Size is the natural logarithm of a firm’s asset value. BTM is the book-to-market ratio. RET is a firm’s gross market returns over the past 12 months and RETVOL stands for the standard deviation in each firm’s monthly returns over the past 12 months. CAPEX is the lagged capital expenditure as a percentage of total assets and ROA is the lagged return on assets. LOSS equals 1 for firms that report a loss in any of the previous four quarters and 0 otherwise. ISSUE equals 1 if the firm issues equity in that quarter and 0 otherwise. Coverage is the number of analysts following a firm in a certain quarter. IOHLD is the institutional ownership. Panel B reports the regression results for the baseline model in Equation (5). Panel C reports the probit regression result for the first-stage of the Heckman two-stage procedure. Panel D reports the regression result for Equation (7). Panel E presents the reaction of stock prices to different types of stock recommendations on local and nonlocal firms, respectively.

Panel A: Summary statistics

Variable	N	mean	sd	p25	p50	p75
Strongbuy	81738	3.16	8.46	0.00	0.00	3.00
Buy	81738	4.45	9.17	0.00	0.00	4.00
Holdorworse	81738	0.89	2.80	0.00	0.00	0.00
Coverage	81738	2.03	3.67	0.00	0.00	2.00
IOHLD	81738	2.28	4.91	0.00	0.01	1.98
BTM	81738	0.98	0.84	0.43	0.72	1.24
ROA	81738	0.02	0.04	0.00	0.02	0.04
CAPEX	81738	0.06	0.06	0.02	0.04	0.09
RETVOL	81738	0.12	0.06	0.08	0.11	0.15
RET	81738	0.21	0.54	-0.16	0.08	0.51
Size	81738	21.98	1.04	21.26	21.87	22.60
ISSUE	81738	0.01	0.08	0.00	0.00	0.00
LOSS	81738	0.25	0.43	0.00	0.00	0.00
PASTMARKET	81738	0.13	0.41	-0.11	0.03	0.38
MARKETVOL	81738	0.07	0.03	0.05	0.06	0.08
GDPCHG	81738	0.03	0.01	0.02	0.03	0.04
SENT	81738	-0.02	0.88	-0.68	-0.16	0.38
Local	81738	0.57	0.50	0.00	1.00	1.00
Local*SENT	81738	-0.01	0.66	-0.22	0.00	0.00

Panel B: Regressions analysis for number of stock recommendations and investor sentiment

VARIABLES	Strongbuy	Buy	Holdorworse
SENT	-0.040 (-0.59)	-0.074 (-1.02)	-0.016 (-0.57)
Local	40.930*** (7.17)	26.976*** (6.98)	0.061 (0.06)
Local*SENT	0.114** (2.45)	-0.053 (-1.07)	-0.022 (-1.07)
BTM	0.751*** (15.31)	0.418*** (9.36)	-0.042** (-2.32)
ROA	10.400*** (13.11)	13.624*** (16.27)	-0.323 (-1.10)
CAPEX	1.135*** (2.65)	2.173*** (4.64)	0.811*** (4.42)
Size	3.137*** (43.52)	3.191*** (49.08)	0.446*** (20.20)

IOHLD	0.269*** (27.83)	0.253*** (26.17)	-0.032*** (-11.41)
RETVOL	0.095 (0.20)	1.665*** (3.18)	1.259*** (5.67)
RET	0.729*** (9.00)	0.001 (0.01)	-0.477*** (-13.92)
ISSUE	0.701* (1.90)	0.589* (1.72)	0.081 (0.79)
LOSS	0.084 (1.64)	0.018 (0.32)	-0.081*** (-3.28)
PASTMARKET	-1.469*** (-8.68)	-0.923*** (-5.01)	-0.023 (-0.32)
MARKETVOL	-8.915*** (-4.13)	-17.290*** (-7.26)	0.538 (0.65)
GDPCHG	12.276*** (4.50)	16.407*** (5.05)	-4.084*** (-3.04)
Observations	81,738	81,738	81,738
R-squared	0.474	0.507	0.280
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
Adjusted R-squared	0.460	0.494	0.261

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel C: Heckman two-stage regression (Stage one)

VARIABLES	treated
NUM	0.001*** (7.02)
SENT	0.072*** (5.30)
Local	0.051*** (4.87)
Local*SENT	0.007 (0.57)
BTM	0.145*** (21.81)
ROA	6.374*** (32.13)
CAPEX	2.518***

	(29.15)
Size	0.603***
	(83.97)
IOHLD	0.098***
	(39.77)
RETVOL	-1.209***
	(-10.84)
RET	0.305***
	(17.07)
ISSUE	0.126*
	(1.81)
LOSS	-0.233***
	(-16.81)
PASTMARKET	-0.586***
	(-21.34)
MARKETVOL	2.469***
	(9.65)
GDPCHG	-5.499***
	(-14.40)
Observations	81,738
Pseudo R2	0.308

Robust z-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel D: Heckman two-stage regression (Stage two)

VARIABLES	Strongbuy	Buy	Holdorworse
SENT	0.438***	0.259***	-0.035
	(6.49)	(3.55)	(-1.25)
Local	37.253***	24.410***	0.208
	(6.98)	(6.61)	(0.22)
Local*SENT	0.097**	-0.065	-0.021
	(2.11)	(-1.32)	(-1.04)
BTM	1.741***	1.109***	-0.082***
	(30.03)	(21.58)	(-4.23)
ROA	57.392***	46.421***	-2.209***
	(34.92)	(29.83)	(-4.31)
CAPEX	19.608***	15.066***	0.070
	(28.62)	(22.30)	(0.28)
Size	7.067***	5.933***	0.289***

	(49.16)	(48.24)	(7.39)
IOHLD	0.601***	0.485***	-0.045***
	(39.39)	(35.05)	(-11.12)
RETVOL	-9.426***	-4.980***	1.641***
	(-17.19)	(-8.57)	(6.85)
RET	3.211***	1.733***	-0.576***
	(29.62)	(15.95)	(-14.24)
ISSUE	1.715***	1.296***	0.041
	(4.75)	(3.79)	(0.39)
LOSS	-2.252***	-1.612***	0.013
	(-29.69)	(-21.55)	(0.41)
PASTMARKET	-5.314***	-3.606***	0.131*
	(-25.97)	(-17.21)	(1.66)
MARKETVOL	11.158***	-3.280	-0.268
	(5.07)	(-1.35)	(-0.31)
GDPCHG	-27.105***	-11.079***	-2.504*
	(-9.29)	(-3.27)	(-1.78)
inv mills1	11.423***	7.973***	-0.458***
	(36.45)	(29.29)	(-5.32)
Observations	81,738	81,738	81,738
R-squared	0.504	0.520	0.281
Firm FE	YES	YES	YES
Firm FE	YES	YES	YES
Adjusted R-squared	0.491	0.507	0.262

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Panel E: Market reaction to different types of stock recommendations

Variables	No. of Obs.	Local	Nonlocal	Difference
Strongbuy	52,326	1.73%***	1.49%***	0.24%***
Buy	70,770	1.07%***	0.92%***	0.14%***
Holdorworse	18,013	-0.18%***	-0.31%***	0.13%***

*** p<0.01, ** p<0.05, * p<0.1

Table 15. A comparison of institutional ownership

This table compares the quarterly institutional ownership between firms. Specifically, Panel A reports the comparison of institutional ownership between local and nonlocal firms; Panel B summarizes the comparison of institutional ownership between firms with above- and below-sample median sentiment index and between local and nonlocal firms. 250 miles²⁷ is the geographic cutting boundary for local and nonlocal firms.

Panel A: Comparison of institutional holding on local and nonlocal firms

Geography	
Local	2.33
Nonlocal	2.32
0.40[0.69]	

Panel B: Comparison of institutional holding on local and nonlocal firms, high-sentiment and low-sentiment period.

Geography\Sentiment	Sent≤Sample Median	Sent>Sample Median	t_Value[p_value]
Local	2.00	2.66	Row test: 13.96***[0.00]
Nonlocal	1.95	2.70	Row test: 13.15***[0.00]
Diagonal test: 13.71***[0.00]	Column test: 0.98[0.33]	Column test: 0.60[0.55]	Diagonal test: 13.38***[0.00]

²⁷ The results are robust when I use 100 miles, 100 kilometers as geographic cutting boundary.

Table 16: Gross domestic product and investor sentiment effect

This table reports the results for the effect of GDP on the investor sentiment effect. Panel A presents the regressions of portfolio returns from t to $t+k$ on the sentiment index (SENT), the market earnings-to-price ratio (EP), the risk-free rate (RF), the inflation rate (INFL) and the three Fama–French factors (MKT, SMB and HML) which are shown in Equation (4). I report the coefficient on SENT for each equal-weighted portfolio formed on the nominal value of GDP at the province level. High refers to the average stock returns on firms headquartered in provinces with top ten GDP while Low refers to the returns on firms headquartered in other provinces. High–Low represents the differences in the average stock returns between high GDP and low GDP firms. Panel B shows the same regression results as Panel A but without firms located in Beijing, Shanghai or Guangdong. Simulated p-values are reported in parentheses and are for two-sided tests for the hypothesis that the slope coefficient β does not equal 0.

Panel A: Equal-weighted portfolio regression on high GDP portfolios and low GDP portfolios

Horizons	High	Low	High-Low
9 months	-0.059 (0.23)	-0.061 (0.25)	0.002 (0.64)
12 months	-0.097 (0.06)	-0.100 (0.07)	0.002 (0.55)
15 months	-0.108 (0.03)	-0.109 (0.04)	0.001 (0.65)

Panel B: Equal-weighted portfolio regression on high GDP portfolios and low GDP portfolios (after excluding firms located in Beijing, Shanghai or Guangdong)

Horizons	High	Low	High-Low
9 months	-0.054 (0.28)	-0.056 (0.29)	0.003 (0.57)
12 months	-0.090 (0.08)	-0.093 (0.1)	0.003 (0.58)
15 months	-0.100 (0.05)	-0.099 (0.06)	0.000 (0.81)

Reference

- Aboody, David, Omri Even Tov, Reuven Lehavy, and Brett Trueman, 2016, Overnight returns and firm-specific investor sentiment.
- Agarwal, Sumit, and Robert Hauswald, 2010, Distance and private information in lending, *Review of Financial studies* 23 (7): 2757-2788.
- Almazan, Andres, Adolfo De Motta, Sheridan Titman, and Vahap Uysal, 2010, Financial structure, acquisition opportunities, and firm locations, *The Journal of Finance* 65, 529-563.
- Anand, Amber, Vladimir A Gatchev, Leonardo Madureira, Christo A Pirinsky, and Shane Underwood, 2011, Geographic proximity and price discovery: Evidence from nasdaq, *Journal of Financial Markets* 14, 193-226.
- Antoniou, Constantinos, John A Doukas, and Avanidhar Subrahmanyam, 2015, Investor sentiment, beta, and the cost of equity capital, *Management Science* 62, 347-367.
- Arif, Salman, and Charles MC Lee, 2014, Aggregate investment and investor sentiment, *Review of Financial Studies*, 3241-3279.
- Bagnoli, Mark, Michael B Clement, Michael J Crawley, and Susan G Watts, 2009, The profitability of analysts' stock recommendations: What role does investor sentiment play?
- Baik, Bok, Jun-Koo Kang, and Jin-Mo Kim, 2010, Local institutional investors, information asymmetries, and equity returns, *Journal of Financial Economics* 97, 81-106.
- Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *The Journal of Finance* 61, 1645-1680.
- Baker, Malcolm, and Jeffrey Wurgler, 2007, Investor sentiment in the stock market, *Journal of Economic Perspectives* 21, 129-151.
- Baker, Malcolm, Jeffrey Wurgler, and Yu Yuan, 2012, Global, local, and contagious investor sentiment, *Journal of Financial Economics* 104, 272-287.
- Barber, Brad M, and Terrance Odean, 2008, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *Review of Financial Studies* 21, 785-818.

- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.
- Bergman, Nittai K, and Sugata Roychowdhury, 2008, Investor sentiment and corporate disclosure, *Journal of Accounting Research* 46, 1057-1083.
- Bernile, Gennaro, Shimon Kogan, and Johan Sulaeman, 2015, Information environment and the geography of firms and investors.
- Bernile, Gennaro, Alok Kumar, and Johan Sulaeman, 2015, Home away from home: Geography of information and local investors, *Review of Financial Studies*, 2009-2049.
- Black, Fischer, 1986, Noise, *The Journal of Finance* 41, 528-543.
- Bradshaw, Mark Thomas, 2011, Analysts' forecasts: What do we know after decades of work?
- Brown, Gregory W, and Michael T Cliff, 2005, Investor sentiment and asset valuation, *The Journal of Business* 78, 405-440.
- Brown, Jeffrey R, Zoran Ivković, Paul A Smith, and Scott Weisbenner, 2008, Neighbors matter: Causal community effects and stock market participation, *The Journal of Finance* 63, 1509-1531.
- Brown, Lawrence D, 1997, Analyst forecasting errors: Additional evidence, *Financial Analysts Journal* 81-88.
- Burdekin, Richard CK, and Luke Redfern, 2009, Stock market sentiment and the draining of china's savings deposits, *Economics Letters* 102, 27-29.
- Butler, Alexander W, 2008, Distance still matters: Evidence from municipal bond underwriting, *Review of Financial Studies* 21, 763-784.
- Chang, Eric C, and Sonia ML Wong, 2009, Governance with multiple objectives: Evidence from top executive turnover in china, *Journal of Corporate Finance* 15, 230-244.
- Chen, Ke, and Guo Chen, 2015, The rise of international financial centers in mainland china, *Cities* 47, 10-22.
- Cheng, Qiang, Fei Du, Xin Wang, and Yutao Wang, 2016, Seeing is believing: Analysts'

- corporate site visits, *Review of Accounting Studies* 21, 1245-1286.
- Cornell, Bradford, Wayne R Landsman, and Stephen Stubben, 2014, Accounting information, investor sentiment, and market pricing.
- Corredor, Pilar, Elena Ferrer, and Rafael Santamaria, 2014, Is cognitive bias really present in analyst forecasts? The role of investor sentiment, *International Business Review* 23, 824-837.
- Coval, Joshua D, and Tobias J Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045-2073.
- Coval, Joshua D, and Tobias J Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of political Economy* 109, 811-841.
- Da, Zhi, Joseph Engelberg, and Pengjie Gao, 2015, The sum of all fears investor sentiment and asset prices, *Review of Financial Studies* 28, 1-32.
- De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, 1990, Noise trader risk in financial markets, *Journal of political Economy* 98, 703-738.
- Degryse, Hans, and Steven Ongena, 2005, Distance, lending relationships, and competition, *The Journal of Finance* 60, 231-266.
- Easterwood, John C, and Stacey R Nutt, 1999, Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *The Journal of Finance* 54, 1777-1797.
- El Ghouli, Sadok, Omrane Guedhami, Yang Ni, Jeffrey Pittman, and Samir Saadi, 2013, Does information asymmetry matter to equity pricing? Evidence from firms' geographic location, *Contemporary Accounting Research* 30, 140-181.
- Fama, Eugene F, 1965, The behavior of stock-market prices, *The Journal of Business* 38, 34-105.
- Firth, Michael, Chen Lin, Ping Liu, and Yuhai Xuan, 2013, The client is king: Do mutual fund relationships bias analyst recommendations? *Journal of Accounting Research* 51, 165-200.
- Firth, Michael, Kailong Wang, and Sonia ML Wong, 2014, Corporate transparency and the impact of investor sentiment on stock prices, *Management Science* 61, 1630-1647.

- Friedman, Milton, 1953. *Essays in positive economics* University of Chicago Press.
- Friesen, Geoffrey, and Paul A. Weller, 2006, Quantifying cognitive biases in analyst earnings forecasts, *Journal of Financial Markets* 9, 333-365.
- Garcia, Diego, and Øyvind Norli, 2012, Geographic dispersion and stock returns, *Journal of Financial Economics* 106, 547-565.
- Grinblatt, Mark, and Matti Keloharju, 2001, How distance, language, and culture influence stockholdings and trades, *The Journal of Finance* 56, 1053-1073.
- Han, Bing, and David A. Hirshleifer, 2016, Social transmission bias and active investing.
- Heath, Chip, and Amos Tversky, 1991, Preference and belief: Ambiguity and competence in choice under uncertainty, *Journal of Risk and Uncertainty* 4, 5-28.
- Hong, Harrison, and Jeffrey D Kubik, 2003, Analyzing the analysts: Career concerns and biased earnings forecasts, *The Journal of Finance* 58, 313-351.
- Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2004, Social interaction and stock-market participation, *The Journal of Finance* 59, 137-163.
- Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2005, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, *The Journal of Finance* 60, 2801-2824.
- Hong, Harrison, Jeffrey D Kubik, and Jeremy C Stein, 2008, The only game in town: Stock-price consequences of local bias, *Journal of Financial Economics* 90, 20-37.
- Horst, Ulrich, and Jose A. Scheinkman, 2006, Equilibria in systems of social interactions, *Journal of Economic Theory* 130, 44-77.
- Hribar, Paul, and John McNinnis, 2012, Investor sentiment and analysts' earnings forecast errors, *Management Science* 58, 293-307.
- Huberman, Gur, 2001, Familiarity breeds investment, *Review of Financial Studies* 14, 659-680.
- Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer:

- A nonevent that made stock prices soar, *The Journal of Finance* 56, 387-396.
- Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267-306.
- John, Kose, Anzhela Knyazeva, and Diana Knyazeva, 2011, Does geography matter? Firm location and corporate payout policy, *Journal of Financial Economics* 101, 533-551.
- Kang, JunKoo, and Jin-Mo Kim, 2008, The geography of block acquisitions, *The Journal of Finance* 63, 2817-2858.
- Korniotis, George M, and Alok Kumar, 2013, Do portfolio distortions reflect superior information or psychological biases? *Journal of Financial and Quantitative Analysis* 48, 1-45.
- Kubick, Thomas R, and G Brandon Lockhart, 2016, Proximity to the sec and stock price crash risk, *Financial Management* 45: 341–367.
- Lee, Charles, Andrei Shleifer, and Richard H Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *The Journal of Finance* 46, 75-109.
- Lemmon, Michael L, and Karl V Lins, 2003, Ownership structure, corporate governance, and firm value: Evidence from the east asian financial crisis, *The Journal of Finance* 58, 1445-1468.
- Lemmon, Michael, and Evgenia Portniaguina, 2006, Consumer confidence and asset prices: Some empirical evidence, *Review of Financial Studies* 19, 1499-1529.
- Leuz, Christian, Dhananjay Nanda, and Peter D Wysocki, 2003, Earnings management and investor protection: An international comparison, *Journal of Financial Economics* 69, 505-527.
- Li, Si-ming, and Yi-man Shum. "Impacts of the national trunk highway system on accessibility in China." *Journal of Transport Geography* 9.1 (2001): 39-48.
- Lim, Terence, 2001, Rationality and analysts' forecast bias, *The Journal of Finance* 56, 369-385.
- Lin, Hsiou-wei, and Maureen F McNichols, 1998, Underwriting relationships, analysts' earnings forecasts and investment recommendations, *Journal of Accounting and Economics* 25, 101-127.

- Lin, Yatang, 2016, Travel costs and urban specialization patterns: Evidence from china's high speed railway system, *Journal of Urban Economics* 98, 98-123.
- Loughran, Tim, 2007, Geographic dissemination of information, *Journal of Corporate Finance* 13, 675-694.
- Loughran, Tim, and Paul Schultz, 2005, Liquidity: Urban versus rural firms, *Journal of Financial Economics* 78, 341-374.
- Malloy, Christopher J, 2005, The geography of equity analysis, *The Journal of Finance* 60, 719-755.
- McLean, R David, and Mengxin Zhao, 2014, The business cycle, investor sentiment, and costly external finance, *The Journal of Finance* 69, 1377-1409.
- Mikhail, Michael B, Beverly R Walther, and Richard H Willis, 1999, Does forecast accuracy matter to security analysts? *The Accounting Review* 74, 185-200.
- Mokoaleli-Mokoteli, Thabang, Richard J Taffler, and Vineet Agarwal, 2009, Behavioural bias and conflicts of interest in analyst stock recommendations, *Journal of Business Finance & Accounting* 36, 384-418.
- Neal, Robert, and Simon M Wheatley, 1998, Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis* 33, 523-547.
- Nelson, Charles R, and Myung J Kim, 1993, Predictable stock returns: The role of small sample bias, *The Journal of Finance* 48, 641-661.
- Ng, Lilian, and Fei Wu, 2006, Revealed stock preferences of individual investors: Evidence from chinese equity markets, *Pacific-Basin Finance Journal* 14, 175-192.
- O'Brien, Patricia C, and Hongping Tan, 2015, Geographic proximity and analyst coverage decisions: Evidence from IPOs, *Journal of Accounting and Economics* 59, 41-59.
- Petersen, Mitchell A, 2004, *Information: Hard and soft*.
- Pirinsky, Christo, and Qinghai Wang, 2006, Does corporate headquarters location matter for stock returns? *The Journal of Finance* 61, 1991-2015.
- Pirinsky, Christo A, and Qinghai Wang, 2010, Geographic location and corporate

finance, World Scientific Book Chapters, 23-38.

Qian, Hong, 2009, Time variation in analyst optimism: An investor sentiment explanation, *The Journal of Behavioral Finance* 10, 182-193.

Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987-2010.

Seybert, Nicholas, and Holly I Yang, 2012, The party's over: The role of earnings guidance in resolving sentiment-driven overvaluation, *Management Science* 58, 308-319.

Sharif, Saqib, Hamish D Anderson, and Ben R Marshall, 2014, Against the tide: The commencement of short selling and margin trading in mainland china, *Accounting & Finance* 54, 1319-1355.

Shiller, Robert J, Stanley Fischer, and Benjamin M Friedman, 1984, Stock prices and social dynamics, *Brookings Papers on Economic Activity*, 457-510.

Stambaugh, Robert F, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

Stambaugh, Robert F, Jianfeng Yu, and Yu Yuan, 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* 104, 288-302.

Swaminathan, Bhaskaran, 1996, Time-varying expected small firm returns and closed-end fund discounts, *Review of Financial Studies* 9, 845-887.

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

Yang, Zhifeng, 2013, "Do political connections add value to audit firms? Evidence from IPO audits in China." *Contemporary Accounting Research* 30, 891-921.

Zhu, Ning, 2002, The local bias of individual investors.

Zhao, Simon Xiaobin, Qionghua Lao, and Chan Neo Ying Ming, 2013, The rise of china and the development of financial centres in hong kong, beijing, shanghai, and shenzhen, *Journal of Globalization Studies* 4(1).

Appendix: Variable Definitions

Variable	Definition
CEFD	The value-weighted average difference between the net asset value (NAV) per share and market price of each closed-end fund share, divided by the net asset value (NAV) per share
TURN	The natural log of the ratio of the amount of monthly market turnover trading to the aggregate market value, detrended by the 5-month moving average
NIPO	The number of monthly initial public offerings
RIPO	The average market-adjusted first-day returns of initial public offerings
Eshare	The ratio of equity issues to total equity issues and bank borrowing.
DSG	The growth rate of savings deposits controlling for the growth of money supply (M0)
NACT	The natural log number of new investment accounts (in thousands)
IPG	The growth in industrial production
CG	The consumption growth rate
sentiment	The sentiment index
EP	The market level earnings-to-price ratio
RF	The risk-free rate, measured by the monthly interest rate on one-year fixed deposits
INFL	The monthly inflation rate
ExRET	The excess return on the market portfolio of stocks
SMB	The difference between average returns on the small-stock portfolio and average returns on the large-stock portfolio per month
HML	The difference between average returns on the high-BE/ME portfolio and average returns on the low-BE/ME portfolio per month
cr_eq_f <i>i</i>	The <i>i</i> -month future market return
Local	The average returns of firms located within certain radius (250miles, 100miles or 100kilometers) of Beijing, Shanghai or Shenzhen in Table 5 to Table 10 and in Table 15
Local	The average returns on firms headquartered in cities with less travel times from Beijing, Shanghai or Shenzhen in Table 11
Local	A dummy variable that equals 1 if the firm is identified as a local firm and 0 otherwise in Table 14
Nonlocal	The average returns of firms located outside certain radius (250miles, 100miles or 100kilometers) of Beijing, Shanghai or Shenzhen in Table 5 to Table 10 and in Table 15
Nonlocal	The average returns on firms headquartered in cities with more travel times from Beijing, Shanghai or Shenzhen in Table 11
D1	The average returns of firms located within a 100-kilometer radius of Beijing, Shanghai or Shenzhen

D2	The average returns of firms located between 100-kilometer radius of Beijing, Shanghai or Shenzhen and 1,000-kilometer radius of Beijing, Shanghai and Shenzhen
D3	The average returns of firms located outside a 1,000-kilometer radius of Beijing, Shanghai and Shenzhen
High	The average returns on firms located in provinces with top three stock market participation rate/top ten GDP
Low	The average returns on firms located in provinces with low (lower than top three) stock market participation rate/ low GDP (lower than top ten)
MPR	The provincial-level total stock accounts divided by the year-end population in that province
Strongbuy	The number of quarterly “Strong buy” recommendations
Buy	The number of quarterly “Buy” recommendations
Holdorworse	The sum of quarterly “Netural”, “Hold” and “Sell” recommendations
Coverage	The number of analysts following a firm in a quarter
IOHLD	The institutional ownership in a quarter
BTM	The book-to-market ratio
ROA	The lagged return on assets
CAPEX	The lagged capital expenditure as a percentage of total assets
RETVOL	The standard deviation in each firm’s monthly returns over the past 12 months
RET	A firm’s gross market returns over the past 12 months
Size	The natural logarithm of a firm’s asset value
ISSUE	A dummy variable that equals 1 if the firm issues equity in that quarter and 0 otherwise
LOSS	A dummy variable that equals 1 for firms that report a loss in any of the previous four quarters and 0 otherwise
PASTMARKET	The gross returns on the value-weighted market index over the 12 months prior to the beginning of a quarter
MARKETVOL	The volatility of gross market returns during the 12 months prior to the beginning of a quarter
GDPCHG	The seasonally adjusted percentage change in GDP in a quarter
SENT	The average monthly investor sentiment index in a quarter

---