Master Thesis, Spring 2020
Department of Finance
Stockholm School of Economics

Does Noise Trader Risk Repel Arbitrageurs?

Evidence from Chinese A-H Share Premia

Haihui Li 41506@student.hhs.se
Qing Zhu 41251@student.hhs.se
Supervisor: Jungsuk Han, Associate Professor

Contents

Abstract

What causes Chinese A-H share premia puzzle? A-shares enjoy a premium over corresponding H-shares on average by 125%, despite the same rights and dividends. Existing hypotheses such as differential risk, differential demand, liquidity, and asymmetric information cannot successfully account for the great magnitude of inflated A-share prices and are also inconsistent with our sample from 2014-2019. In this paper, we propose and empirically test a simple new explanation of A-H premia based on the combination of noise trader risk and the limits to arbitrage. We argue that arbitrageurs avoid correcting mispricing because of expected future price volatilities driven by investor sentiment. Our methodology can explain the magnitude of A-H premia as well as the evolution of A-H premia using empirical proxies for noise trader risk.

Key words

noise trader risk, A-H premia, volatility, investor sentiment, limits to arbitrage.

Acknowledgement

We would like to extend our great gratitude towards Associate Professor *Jungsuk Han* for the patient guidance. From the very beginning, his in-depth knowledge about the pricing premia, current market and previous literatures have consistently led us to every next possible step. He kept encouraging us to adopt more creative approaches and to go through all the difficulties even when the road got tough. Without his timely advice, the goal of the thesis would not have been met. Once again, we deliver our sincere gratitude towards *Jungsuk Han*.

1. Introduction

A-H share premia has long been an intriguing puzzle in China's stock market, since it is brought up by Baily (1994) and Worldbank (1997). A-shares are domestic shares, issued and traded only in mainland China A-share market, and H-shares are foreign shares issued and traded in Hong Kong's stock market. However, the puzzle is that for cross-listing companies which issue both A-shares and H-shares, price of an A-share always enjoy a premium over that of the corresponding H-share stock, despite that both types of shares have basically the same voting right and dividends. An extreme case is that by April 2020, thirteen out of ninety-three cross-listing A-H companies enjoyed A-H premia by more than 200%, among which Luoyang Glass reached an astounding premium at 573.17%.

Previous literatures on dual-listing pricing disparity phenomena suggest market segmentation as a main explanatory standpoint. In general, existing papers bring up four different hypotheses to address the puzzle of dual listing share disparity: liquidity hypothesis (Poon et al, 1998), asymmetric information hypothesis (Chakravarty et al, 1998), differential risk hypothesis (Sun and Tong, 1999), and differential demand hypothesis (Stulz and Wasserfallen, 1995). These hypotheses suggest that the premia of A-shares over their H-share counterparts can be explained by different risk premia and risk differences in the segmented markets. However, these hypotheses cannot successfully justify the great magnitude of mispricing in the aforementioned phenomena, for example, a 573.17% A-H premium. Our examination of these hypotheses with a panel data sample in the period from 2014 to 2019 reveals that only 57% of the variation of A-H premia can be explained by them[1]. Further, our empirical results are inconsistent with implications of these hypotheses. Therefore, a different explanation is needed.

In this paper, we propose and empirically test a simple new explanation of A-H premia based on the combination of investor sentiment and the limits to arbitrage. We argue that arbitrageurs avoid correcting mispricing because of expected future price volatilities arising from investor sentiment. We find the following empirical results

[1] See Appendix c. Original Regression Results

which support our main argument. First, we find that A-H premia is significantly correlated with the measure of optimistic investor sentiment. Second, we find that A-H premia has a significant predictive power of future stock market volatility. This indicates that arbitrageurs avoid correcting mispricing because of expected future price volatilities. Third, we find that the correlation between A-H premia and investor sentiment becomes stronger during the period with higher limits to arbitrage.

According to the efficient market hypothesis, ideally, if arbitrageurs can find perfect substitutes of asset, they can arbitrage and correct mispricing. However, even if they can find perfect substitutes, they might still fail to correct mispricing due to noise trader risk. According to Black (1986), noise traders are investors who irrationally believe that the noise they have received gives them advantages thus will act on such noise. When there is a surge of investor sentiment, noise traders will diverge stock prices from their fundamental value. In an efficient market, arbitrageurs will step right in and trade in the opposite direction against noise traders, taking positions as largely as possible, and thus the mispricing will be corrected. However, De Long et al (1990) suggest that the uncertainty of noise traders' future beliefs, which may diverge the prices even further from the fundamental value against arbitrageurs, impose a noise trader risk on arbitrageurs, therefore arbitrageurs will become less willing to bet against noise traders due to risk aversion. For instance, if a surge of optimism in the market drives the share price higher than the fundamental value, arbitrageurs will become more conservative and less willing to sell short against unsophisticated noise traders, fearing that the noise traders may get even more fervent and drive the price even higher in the future. As a result, stock prices can either be inflated or deflated by noise traders.

If noise traders' beliefs are randomly distributed, their trading won't affect the price, as their noise trading cancel out each other. However, related studies suggest against this randomness assumption. Shiller (1984) shows that inexperienced investors make trading decisions based on their beliefs, and their trading are correlated with each other. In China's A-share market, Zhou (2007) examines that both A-share market and B-share market have herding behavior. Tan et al (2008) show herding behavior of AB-shares and A-share investors have higher herding tendency than B-share investors: when market is rising, A-share investors have higher trading volumes and volatilities than B-share investors. If noise traders' beliefs tend to be optimistic most of the time, it'd be reasonable to argue that noise traders inflate stock prices in a market. As Mei et

al (2005) indicates, strong short-sales constraints trigger speculative motive in China's A-share market. Combining these pieces of jigsaw, we arrive at our explanation for A-H premia and for premia in similar segmented markets.

Our argument is reasonable in the segmented A-share and H-share markets for two reasons. First, the A-share market conditions tend to ferment speculative atmosphere. Strict short sell constraints (Gu et al, 2018) make great difficulties for short-selling, and investors would expect stock prices to go up when arbitrageurs are tie-handed to some extent. In addition, government censorship which holds negative news in control helps to inflate A-share price prices (Dong et al, 2018). Also, China's A-share market currently is dominated by individual investors, who holds 62% of total asset float of the market compared to 20% in the US, and their beliefs undoubtedly is a strong drive of the market. These individual investors are inexperienced in a sense that China has opened its modern stock market for less than 30 years. Second, the corresponding H-shares are ideal control for asset fundamental value, through which we can separate the mispricing portion in the form of A-H premia from the fundamental value.

We use an unbalanced panel data of dual-listing A-shares and H-shares in the period from 2014 to 2019 to examine our hypothesis. We first examine whether optimistic beliefs create noise trader risk and inflate A-share prices. We then examine whether high A-H premia foretell high future market volatility in the A-share market, a risk perceived by arbitrages. We also include an examination of the effect of noise trader risk on A-share price inflation. The empirical results support our argument.

First, the optimistic sentiment effect on A-H premia is positive and significant, while the pessimistic sentiment effect is insignificant, which shows that pessimistic sentiment is mitigated. Second, the results show that high A-H premia foretell high market volatility in the future, which persists for over one year. It suggests arbitrageurs can read from a high level of A-H premia that there are high risks involved. Third, the optimistic sentiment effect on the future market volatility is significant and positive and lasts for over one year. The pattern of the optimistic sentiment effect is similar to that of premium on market volatility. It suggests that noise trader risk is a source of future market volatility. On top of that, the effect of future market volatility predicted by optimistic sentiment has a significant and positive impact on A-H premia, which shows

that expected volatility inflates the price. Fourth, our regression results show that the predicted future market volatility has a positive and significant coefficient on A-H premia. It reveals that noise trader risk inflate A-share prices.

These series of empirical results close the logic circuit of our explanation: A-H premia capture future market volatility: higher A-H premia imply higher future market volatility. When the price is highly inflated, arbitrageurs are repelled by the future market volatility, allowing assets to be inflated in the market. One source of the future market volatility is noise trader risk triggered by enthusiastic noise traders in a market with strict short-sales constraints.

Our research enriches the literature of A-H premia by offering a new explanation for A-H premia. To our knowledge, we are the first ones to adopt noise trader risk hypothesis to explain for A-H premia and empirically test the hypothesis with A-H premia data. This paper not only helps to combine the competing hypotheses, which focus on differential risk premia and different risks received (Fernald and Rogers, 1999; Li et al, 2006; Guo et al, 2013), from the inefficient market point of view, but also illustrates a dynamic explanation for A-H premia compared to the static ones. We also answer the questions unaddressed by Mei et al: how investor sentiments cause inflated prices.

This thesis is divided into nine sections. In section one, we introduce our research and brief on our findings. In section two, we review previous literature and indicate the location of our study. In section three, we cast a glimpse on the market background, by introducing China's stock market and the A-H premia puzzle. In section four, we list out our assumptions, develop, and elaborate on our theory. In section five, we demonstrate our empirical methodology and report our data sources. In section six, we examine the empirical results. Section seven introduces alternative hypotheses. In section eight, we complete our findings with supplementary empirical results and discusses policy implications and investment recommendation. We conclude our thesis in section nine.

2. *Literature Review*

The puzzle of Chinese specific A-share to foreign share premium is firstly brought by Baily (1994) and Worldbank (1997). In general, there are two branches of study attempting to explain for A-H and A-B premia -- market segmentation and inefficient market, which represent the rational portion and the irrational portion of A-H premia respectively.

From the market segmentation point of view, four different hypotheses are proposed to explain A-B-share premia and A-H share premia, including: differential risk hypothesis (Eun and Janakiramanan, 1986), differential demand hypothesis (Stulz and Wasserfallen, 1995), liquidity hypotheses (Amihud and Mendelson, 1986; Wruck, 1989; Silber, 1991), and asymmetric information hypothesis (Bailey and Jagtiani, 1994). According to these theories, segmentation separates two or several groups of investors into different markets, both geographically and systematically. Investors of different groups value assets differently for three major reasons: investors' required rate of return for taking the same risk differ, and the magnitude of risk-averseness vary toward the same amount of risks; different market indeed cast distinctive risk levels to investors therein; investment universe, i.e., financial instrument options vary in different market in terms of market development and trading restrictions imposed on individual investors. A common approach to explain such hypothesis is by pricing in different hypothesis factors in terms of risk premia, and the empirical results show that most of the times aforementioned mispricing can be pinned down to market segmentation theory. In section seven, we discuss the details of the literature on these four hypotheses.

In combination with all these four hypotheses, Fan and Wang (2016) find that there is empirical evidence supporting differential risk hypothesis and liquidity hypothesis in A-H premia with a data sample from January 2013 to December 2015. However, this series of hypotheses cannot successfully explain the magnitude of mispricing. Further, they are not consistent with our data in the sample period. Our empirical results are presented in section eight.

Another standpoint for A-H premia explanation is inefficient market. Mei et al (2005) examine the existence of speculative motive in A-B-share premia. According to them, the limit to arbitrage in the China's stock market stimulates speculative motive,

as the marginal buyers are more optimistic than the current shareholders. In fact, Sun and Tong (1999) suggests that domestic investors are more optimistic than foreign investors about expected growth and the optimistic attitude is a reason for A-B-share premia. However, this line of research only shows that the likes of speculative motive exist in the premia, but it remains to be elaborated on how the investor beliefs affect the premia, or why the premia are affected by investor sentiments.

To sum up, former study empirically examine that A-H premia can be partially explained by differential risk hypothesis, liquidity hypothesis, and speculative motive in the A-share market. However, as mentioned earlier, the market segmentation hypotheses in the previous literature cannot successfully explain the magnitude of mispricing. Further, they are not consistent with our data in the sample period. In addition, the inefficient market study only shows the effect of investor sentiments but lacking a dynamic and detailed elaboration on how A-H-share premia evolves with investor beliefs. However, our hypothesis based on noise trader risk can successfully explain the large magnitude of mispricing and answer the question unaddressed by previous inefficient market literature. To our knowledge, this is the first paper which tests the hypotheses on A-H premia. Further, we can also explain the evolution of A-H premia using empirical proxies for noise trader risk.

3. Market Background

3.1 Characteristic of China's Stock Market

In 1990, the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE) established. Their establishment marks the beginning of China's modern secondary market. Only 8 equities and 22 bonds were listed on the SSE in the first year. As of April 14th, 2020, there were in total 3736 companies listed on the Shanghai and Shenzhen stock exchanges, with total market capitalization of 57.279 trillion RMB, among which 47.377 trillion RMB are free float.

The Chinese stock market has its unique characteristics, even though its capital market structure and its legal and financial systems have been developing and learning from more developed markets.

First, China's stock market has a strong financing assistance purpose for the state-run sectors of the economy (Allen and Shen, 2012). The dominant presence of state-run companies in China's stock market is definitely unneglectable. As of April 2020, 30% of the listed companies in the A-share market are state-run, which is directly under control of local or central government. These companies occupy 85% of total assets, 54% of total market capitalization, 73% of total revenue, and 72% of total dividend in the market.

Second, the nature of investors is highly distinguishable. Individual investors, rather than institutional ones, show higher participation rates than investors in more developed markets. As of 2019, Chines individual investors hold 62% of the total asset float (the total value of outstanding shares), whereas 20%, 13.5% and 17% are for US, UK and Japanese market respectively. Even though the ratio is decreasing, at a rather slow pace, individual investor behavioral patterns are still one of the confirming main drivers for stock pricing fluctuations in Chinese capital market.

Third, both its listing and delisting mechanism are under the government's control. The initial public offering process (IPO) in China is under an approval-based IPO system, where approval of the China Securities Regulatory Commission (CSRC) is required and is usually longer and more costly than ones in the US (Liu, Stambaugh, and Yuan, 2019). However, both delisting and listing regulations and constraints are

not yet perfect, and listed companies are rarely delisted. The imperfection, as a result, allow for a space for easier quasi-legal ways for going public-back door listing. Listed companies have a special layer of value, for it can be a quicker way to get listed than going through the formal IPO procedure. This value stimulates market speculation on problematic companies, whose stock prices will skyrocket if the companies are used as a shell for back door listing.

Forth, Chinese stock market adopt T+1 settlement rule and $\pm 10\%$ ceilings for daily trading. They are designed to suppress irrational speculative behaviors and unnecessary fluctuations, notwithstanding the disputable real effects.

Fifth, the A-share market has strict short-sales constraints and limit to arbitrage. At first hand, the threshold of trading options and futures are high for individual investors. Individual investors are allowed for shorting and taking part in futures/options transactions only if they hold a minimum balance of RMB 0.5 million for 20 consecutive days in their margin accounts, right before the application for transactions. They should also have a prior minimum 6 months of margin trading/short selling experience. On the other hand, there are only few instruments for short selling.

Sixth, there is price discrepancy between different share classes, even though the shares contain basically the same rights and same cash flows. Domestic A-shares enjoy a premium over foreign B-shares and H-shares, opposite to more developed markets, such as Switzerland, in which domestic shares provide a discount to foreign shares (Stulz and Wasserfallen 1995). In order to fund state-run companies, the authorities allowed companies to issue shares of different classes to domestic investors and foreign investors. Domestic investors trade A-shares, and the latter are allowed to trade B-shares with US dollars. Foreign investors trade H-shares and B-shares. Share class difference depend on both listed exchanges (mainland and Hong Kong), main investor target and issuing currencies. A-share is issued by companies registered and listed in Chinese mainland, priced and traded in Chinese RMB, foreign ownership upper limit as 30%, which is far from controlling interest requirement of 51%. B-share is issued by corporations registered and listed in mainland, priced in RMB but traded in foreign currencies. B-share originally was only allowed for foreign investors, but after 19th Feb 2001, the CSRC has loosened its entrance constraint for domestic investors and is allowing qualified identities to participate in B-share transactions. However, the upper

limit of US dollar for Chinese citizens is 50,000 per year. Nevertheless, the market cap of total B-share amount is less than 0.1%, as investors from and outside of mainland China has more and more variable channels to access Chinese capital for portfolio diversification thanks to gradual ongoing opening up policies from Chinese government, which in turn diluted investors interest toward B-share significantly. H-share is issued by companies registered in mainland but trade in Hong Kong, thus facing broader range of investors both in term of geography and legal entity classes and is confronted with less trading constraints such as shorting and derivative transactions. However, A-shares consistently enjoy a premium over B-shares and H-shares.

These features characterize China's stock market. Limits to arbitrage and the dominance of individual investors in the secondary market may ferment speculative atmosphere.

3.2 Development of China's Stock Market

Apart from its characteristics, the development of China's stock market is special, for it's under close control of related authorities. China's stock market is becoming more open to investors from outside mainland China, compared to its previous near total restrictions, even towards investors from Hong Kong, against foreign capital injections. In 2002, China launched the Qualified Foreign Institutional Investors (QFII) scheme, in an attempt to bring qualified foreign investors into the A-share market. Qualified foreign investors must fulfill a set of criteria and are under numerous restrictions such as daily capital trading limit and lock-up period. In 2005, the split-structure reform was launched, which aimed to free up non-tradable shares and convert them into tradable shares. The initiative was carried out through negotiations between non-tradable share owners and investors, rapidly boosted the number of outstanding shares but at the same time suppressed the stock prices. In order to attract overseas investors to satisfy the whole economy's growth need, QFII in 2018 removed lock-up period restrictions and some other constraints. Similarly, RMB Qualified Foreign Institutional Investor (RQFII) was established in 2011. RQFII regulates securities traded over RMB whereas QFII over foreign currencies.

Meanwhile, domestic investors are gradually allowed to make investments outside the domestic market. Qualified Domestic Institutional Investors (QDII) scheme,

which was granted in 2006, allows qualified domestic institutions to invest in overseas markets up to 8% of their total assets. In 2014, Shanghai Hong-Kong Stock Connect program was launched, which allows domestic investors to invest in selected stocked in Hong Kong Stock Exchange. This program also provides a channel for investors in Hong Kong to invest in the A-share market.

3.3 Parallel Market and A-share Premia

A-shares consistently enjoy premia over corresponding B-shares and H-shares, despite that dual-listed companies issuing both A-share and B/H-share. Majority of them, nowadays, grant the same voting and dividend rights to both B/H-share classes regardless of geographical diversion of their listed exchanges. B-shares are traded at discounts of 60 - 80% to A-shares on average (Carpenter et al, 2017). H-shares are at discounts of 125% to A-shares on average according to our calculation. B-share market has been suffering from its illiquidity. Before the B-share market was open to domestic investors in 2001, A-B premia reached 80% on average[2], but the price gap narrowed after the policy was implemented. A-H premia also goes through different stages, which are mentioned in later section about policy shifts. Darrat and Zhong (2010) investigate the reason for persistent B-share discount even after lifting restriction toward domestic investors in 2001. They find that long-term equilibrium holds for the relationships between A-share and B-share, and the relationship strengthened in the post-lifting period. However, at the same time, the result ruled out the information asymmetry theory which goes opposite to other conclusions such as Jing and Eddies (2007), which suggest that the stricter corporate disclosure, which is implemented in foreign market as an effective method to lessen the political risks stemming from ownership structure, may lead toward accounting manipulation and transparent disclosure.

Previous research suggests that investors in the B-share and H-share market are more rational than in the A-share market. Wei Sun (2014) points out four possible explanations: foreign investors' less restrictions on information barriers, more alternatives in terms of financial instruments such as shorting, leveraging and derivatives, stricter accounting and disclosure rules and more percentage of institutional

[2] Choice Finance Terminal Data

investors participation ratio.

In addition, B-share and H-share are more integrated with world financial market. Zeng and Luo (2009), applying GJR-GARCH-DCC model, test conditional correlation among A, B and H-share market upon data between 1995 and 2008. They find dynamic and asymmetric correlation between the A and B-share along with B and H-share market, and dynamic but not asymmetric correlation between A and H-share market. They also find increasingly integrated relationships of both B and H-share market with A-share market. Ting and Sie (2005) find that substitution effect from B-shares cross-listed in both US and China. They find significant negative correlation between the B-share premium and the trading volume of Chinese firms traded in the U.S. The effect is stronger than Chinese stocks listed in Hong Kong. In addition, home market price volatility decreases upon the dual listing, but impose no obvious impact to home market liquidity.

3.4 Policy Shift and A-H Premia

In general, A-H premia have gone through different stages. The Hang Seng A-H Premia Index displays the evolution. At the very beginning, H-shares had premia over A-shares before 2006. This period was before the launch of QDII and closely after the split-structure reform. After the implementation of QDII, A-H premia increased and priced that, at a time, A-shares were priced twice as much as H-shares in 2008. However, A-H premia sharply decreased afterwards. During the period around 2011, A-H premia reached a relative bottom and remained low at around 100% until 2015. In 2015, A-H premia started to increase drastically and reached a stable plateau above 120%. During this period, two important policies or events took place: Shanghai Hong-Kong Stock Connect program and H-share Full Convertibility Project.

Formally launched on 17th November 2014, Shanghai Hong-Kong Stock Connect program allows investors from Mainland China and Hong Kong to trade eligible shares on each other's market via brokers. Qualified Chinese investors are able to purchase H-shares on HKEx through domestic securities companies in RMB, and the mechanism is the other way around for Hong Kong investors, in HKD. Government designated specific brokerages in both stock exchanges to implement the trade. Investors, nevertheless, were still under several restrictions. Mainland Chinese

investors faced minimum account balance (RMB 0,5 million) threshold. Only certain shares were eligible for trading, including 568 A-shares and 264 H-shares, in 2014. The A-shares are the dual-listing shares and the component shares of SSE 180 index and SSE 380 index. The H-shares are the dual-listing shares and the component shares HSI index and HSMI index. Moreover, there were total daily quotas for each channel, Shanghai Stock Channel higher than the Hong Kong Stock Channel, RMB 13 billion and RMB 10.5 billion respectively[3]. On 1st May 2018, the quotas increased by three-folds, Shanghai to RMB 52 billion, Hong Kong to RMB 42 billion. Quotas of QDII and QFII also increased at around the same time. Most of the time, inflows from Hong Kong to Mainland are higher than the other direction, and higher volatility in HKEX to SSE channel is observable after quota boost.

H-share Full Convertibility Project was announced on 29th December 2017, by China Securities Regulatory Commission (CSRC), which allows domestic non-tradable H-share owners to convert shares on hand into tradable H-shares. However, the announcement didn't define crystal clearly potential eligible H-share categories since H-share companies can be divided into two types: complete H-share companies and dual-listing companies.

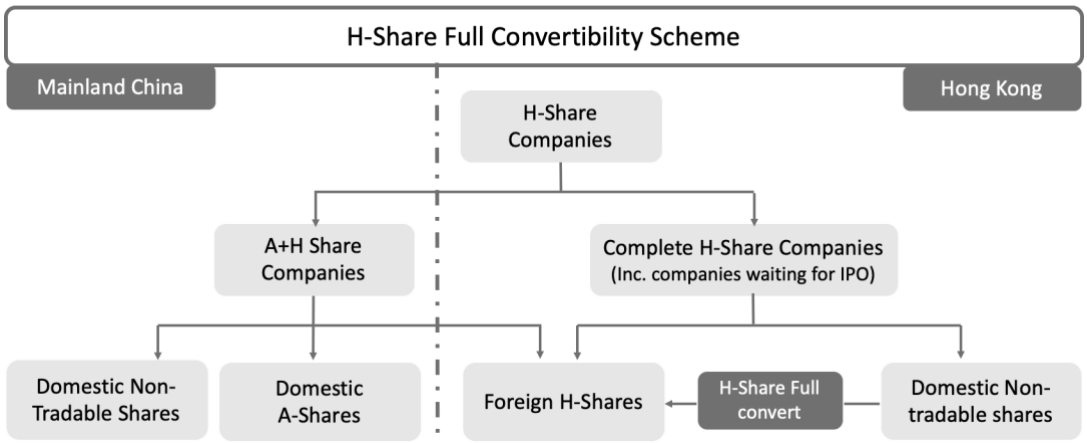

Figure 3.1 H-share Full Convertibility Scheme

H-share Full Convertibility Project allows domestic non-tradable H-share owners to convert shares on hand into tradable H-shares.

[3] China Securities Regulatory Commission document

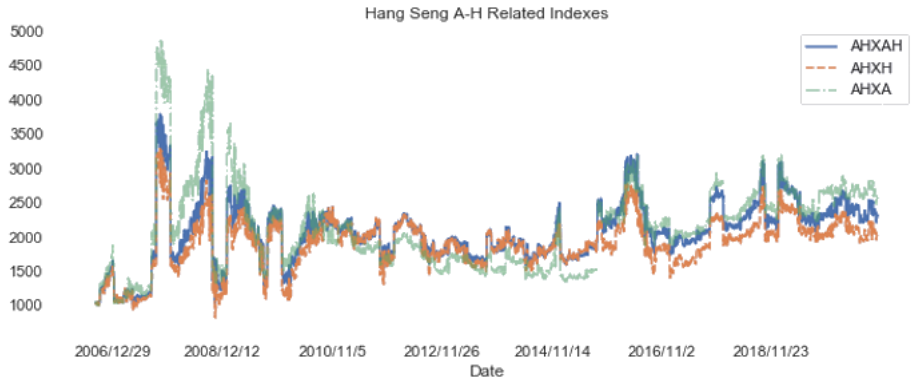

Figure 3.2 Hang Seng A-H Related Indexes

AHXAH: Hang Seng Stock Connect China AH (A+H) Index; AHXA: Hang Seng Stock Connect China AH (A) Index, AHXH: Hang Seng Stock Connect China AH (H)

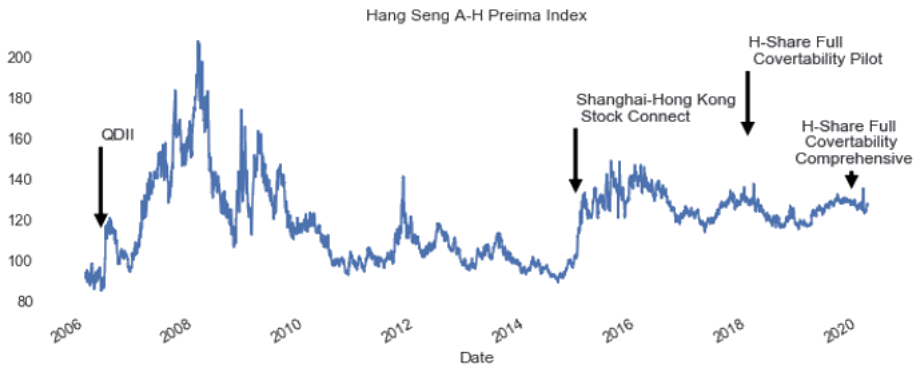

Figure 3.3 Hang Seng A-H Premia Index with major events

The announcement indicated that the scheme follows the one-at-a-time principle and no more than three H-share companies in total are allowed to implement the conversion. H-share companies have full discretionary over the percentage of equity shares for conversion. Three pilot companies were screened by CSRC initially: Legend Holdings Corporation (03396.HK), AVIC Jonhon Optronic Technology (02357.HK), and Weigao Group (01066.HK). Legend Holdings Corporation conducted a partial conversion, led to 46% domestic shares (both tradable and non-tradable) and 54% H-shares of total equity. AVIC Jonhon Optronic Technology and Weigao Group conducted full conversion, resulted in 100% H-shares equities. Upon the completion of conversion, all three companies' stock price increased. On 16th November 2019, all-round implementation of the convertibility scheme kicked off, still following one-at-a-time principle. Thus both H-shares companies and companies waiting to be listed via Hong Kong Stock Exchange were qualified -- about 160 listed H-share companies, the non-tradable shares of which were worth of 5.1% of the total market capital of the Hong Kong Stock Exchange, and awaiting companies estimated to be worth of 7.1% of the total market capital of the Hong Kong Stock Exchange. Nonetheless, dual-listing companies are not included in this scheme.

Cai, McGuinness, and Zhang (2011) suggest that policy and corporate governance changes are principal forces that drive the efficiency of A-H-share premia in three aspects: the long-term expectation of the H- (to A-price) discount, the short-term co-movements in A and H-share prices, and the magnitude of error corrections. Hou and Lee (2014) examine the impact of: non-tradable shares reform in 2005 and finds that the reform significantly contributed to the foreign share discount among state-run firms. It implies that the reform may benefit the minority shareholders by aligning state and private shareholders to monitor managers. Yuan, Zhou, and Li (2018) examine the QDII policy, the QFII policy, and the Shanghai–Hong Kong Stock Connect policy and find that QDII policy has significantly improved the integration of pricing dynamics of A and H-shares. Fan and Wang (2017) suggests that the Shanghai–Hong Kong Stock Connect policy helps reduce A-H premia. In spite of the series of reform, A-H-share premia still and the H-Share Full Convertibility policy in 2018 remains untested.

4. Hypothesis

Our explanation for part of A-H premia is based on De Long et. al.'s noise trader risk study. In this section, we first list out the assumptions and then elaborate the hypothesis in the context of A-share and H-share market.

Our first assumption is that the stock markets, including A-share and H-share market, consist of three groups of investors: noise traders, speculators, and passive investors. This assumption covers different types of investors in stock markets. In a short-time horizon, we deem the composition of these three groups as fixed, since it is hard and costly to have a dramatic change in one's investment strategy. Empirical evidence suggest that noise traders exist in China' A-share market: Zhou (2007) examines that both A-share market and B-share market have 'herding' behavior. Tan et al (2008) show herding behavior of AB-shares and A-share investors have higher herding tendency than B-share investors: when market is rising, A-share investors have higher trading volumes and volatilities than B-share investors. It's also reasonable to assume that arbitrageurs and rational speculators participate in China's A-share market. Chun sheng et al (2005) suggest that irrational investors and limits to arbitrage in China A-share market encourages trading-based stock manipulation.

Noise traders are investors who base their trading decisions on misbeliefs about future assets returns. They form these misconceptions through different channels. For example, a noise trader may blindly trust a broker who claims to have received insider information; an inexperienced investor may learn from his or her friends that the government is going to release a policy beneficial to a certain industry, and so on. Noise traders feel that this noise can yield than excessive returns and trade accordingly. Upon optimistic beliefs, they buy shares and expect the price to rise in the future; when they have pessimistic beliefs, they sell shares with expectation that the price is going to decrease in the future within their investment horizon. If noise traders have randomly distributed beliefs, however, their trading behavior may not affect stock prices because their positions cancel out each other. However, Shiller (1984) finds that inexperienced investors who act on their beliefs have strong correlation in their trading activities. In other words, noise traders tend to have similar beliefs in a stock market. By noise traders, we refer not only to inexperienced individual investors. Under certain performance

pressure, sophisticated institutional investors may also ride on the noise. Lakonishok et al (1991) finds evidence that pension fund managers tend to oversell poorly performing stocks to "window-dress" their performance.

Speculators include both rational speculators and arbitrageurs. Rational speculators predict the act of noise traders, trade beforehand in the same direction, ride on the bandwagon, and exit the position by taking the opposite direction towards the peak. As De Long et al (1990) demonstrate that rational speculators like George Soros destabilize prices. In other words, they diverge the price from the fundamental value in the same direction as noise traders do; even though they try to profit by trading in the opposite direction, the power of noise traders dominates the market. Arbitrageurs are investors who learn the fundamental value and bet actively against the mispricing by taking long positions on underpriced assets or selling shorts of overpriced shares.

Passive investors only make their trading decisions based on fundamental value and are indifferent towards investor sentiments. Passive investors' perceptions of fundamental value are based on expected future cash flows and risk premia, which are associative with differential risk premia hypothesis, liquidity hypothesis, etc. Existing literatures about market segmentation can be applied to passive investors, and a portion of A-H premia can be explained through passive investors. For instance, H-share investors require a discount for illiquidity risk of H-shares, while corresponding A-share investors don't require such a discount. In this way, A-share prices can be higher than their H-share counterparts.

In our second assumption, we assume that arbitrageurs can find perfect substitute assets and carry out arbitrage without basic arbitrage risk, though perfect substitutes are somehow difficult to find and costly to trade in reality. Basic arbitrage risk is not our concern here, for one can always argue that basic arbitrage risk affects the price correction function of arbitrageurs. This assumption complies with the efficient market requirements. In an efficient market, arbitrageurs can take advantage of mispricing by carrying out trading in the opposite direction, if there is no basic arbitrage risk.

The third assumption is that we consider H-share prices as the fundamental value of the assets. Fundamental value is the future cash flows, discounted back into present value. We use the classic Gordon's Growth Formula to define fundamental value:

$$Fundamental\ Value_t = \frac{Cash\ Flow_{t+1}}{Discount\ Rate - Growth\ Rate} \tag{4.1}$$

The fundamental value is determined by the cash flow (often referred to as dividend) generated by the asset, the growth rate of the cash flow, and discount rate required by the market. We assume that the discount rate is a fair compensation for the total risk embedded in the asset. The fundamental value can only be affected by the profitability of the asset, the growth of the assets, and the total risk embedded in the asset, not by perceptions of investors. This assumption helps to separate mispricing from fundamental value and simplifies the segregation. A-H premia are the unexplained portion of price divided by the rational price or fundamental value. It's demonstrated as following:

$$Share\ Price\ H\ \approx\ Fundamental\ Value \tag{4.2}$$

$$Price\ = Fundamental\ Value + Unexplained \tag{4.3}$$

$$Unexplained = Differnt\ Risk +\ Mispricing \tag{4.4}$$

$$Premia\ = \frac{Share\ Price\ A - Share\ Price\ H}{Share\ Price\ H} \tag{4.5}$$

$$\approx \frac{Unexplained}{Fundamental\ Value} \tag{4.6}$$

When noise traders act on their beliefs and change the price of shares, the fundamental value of shares hasn't been changed, for they don't influence the underlying asset and the discount rate. Likewise, passive investors adjust the prices through the unexplained portion of A-H premia, without affecting the fundamental value of the asset. In this setting, we can see the unexplained portion comoves with A-H premia. In addition, the unexplained portion can be presumably divided into three parts: the portion explained by market segmentation, the portion potentially explained by investor sentiment, and other portions. The portion explained by market segmentation can be associative with differential risk premia, illiquidity discounts, and

so on. It's safe to assume that H-share prices represent the fundamental value. On average A-shares enjoy a premium of over 70%, so we don't miss much by this simplification. Even if H-share prices are inflated, the inflated part is minor even when compared to A-H premia.

In an efficient market, given that perfect substitutes of assets can be found and traded, arbitrageurs take every opportunity to bet against mispricing, as they don't bear any risk in their arbitrages, so mispricing won't last long in the market. However, given the same conditions, there might still be a noise trader risk which makes arbitrageurs more conservative and miss some arbitrage opportunities. Hence, prices of assets can diverge from their fundamental value without being corrected. De Long et al demonstrate with their two-period model that arbitrageurs face the risk caused by noise traders' unpredictable beliefs, so that their aggressiveness of betting against noise traders are dampened, even though there aren't basic arbitrage risk. As a result, the mispricing can hardly be corrected immediately, or it takes a fairly long period of time to correct the price while new mispricing may come up. There is empirical evidence to support the theory of noise trader risk. Sias et al (2001) examine that close-end fund shareholders are compensated for bearing noise trader risk. Scruggs (2007) finds that noise trader risk was high, when the Long-Term Capital Management collapsed in 1998 and the technology bubble went bust in 2000.

In the context of A-H premia, noise traders in the A-share market perceive optimistic/pessimistic beliefs and make the corresponding trade and create a noise trader risk to arbitrageurs. The arbitrageurs are risk adverse and hold back in their arbitraging activities. Thus, the price goes above/below the fundamental value, but it remains to be explained that why A-shares are persistently priced higher than H-shares, or why only optimistic beliefs are pronounced in A-H premia?

Our argument regarding this question is that short-sell constraints in China's stock market mitigate pessimistic beliefs and optimistic beliefs are more influential in noise trading. Previous literatures discuss the effect of limits to arbitrage. Miller (1997) argues that with short-sell constraints, optimists affect the stock prices with their valuation, while investors holding pessimistic views will do nothing instead of selling short. Chen et al (2002) examine Miller's theory with 1979-1998 US's mutual fund holdings data, finding that when the breadths are low (the level of short-sell constraints

is strong), the stock prices are higher. It appears that pessimistic beliefs are mitigated in a market with strong short-sell constraints, and Sun and Tong (2000) argue that domestic investors in China harbor more optimistic expectations than foreign investors. Their study implies that noise traders in China, as part of all participants, tend to be optimistic. Combining with Chen et al's (2001) findings, we may argue that with the existence of strong short-sell constraints, optimistic beliefs are voiced out while pessimistic beliefs are mitigated. Even though there are short-sales constraints, it doesn't mean that arbitrageurs are annihilated completely. They can still arrange sophisticated deals to achieve their goals on a large scale.

With above discussed, we arrive at our explanation for A-H premia: arbitrageurs' aversion to bear noise trader risk in a market with limits to arbitrage gives upward mispricing in A-share prices. Noise traders drive up the prices out of optimistic beliefs, while pessimistic beliefs are mitigated by strong short-sell constraints in China's A-share market. Moreover, optimistic beliefs can reinforce itself through positive-feedback strategies (De Long et al, 1990), as noise traders trade more aggressively when they think their beliefs have been confirmed by the rising stock prices.

5. *Methodology*

In this section we elaborate our models and reason how they examine our theory. This section is divided into three parts. First, we demonstrate our model design and indicate our expected results and how the results support our argument. Second, we elaborate on our data sample and its sources. Third, we supplement our methodology with an event analysis.

We need a set of models to demonstrate our theory. Out of consideration of robustness issue, we use panel data to capture individual effects, which can be omitted variables.

First, we want to examine that optimistic beliefs inflate A-share prices and pessimistic beliefs are mitigated, so we examine the effect of both optimistic and pessimistic beliefs on A-H premia in model (5.1):

$$Premium_{i,t} = Premium_{i,t-1} + Optimistic_t + Pessimistic_t + \varepsilon_{i,t} \qquad (5.1)$$

The notations are: i stands for individual stock, which varies cross-sectionally. t stands for time, which varies over different trading days. $Premium_{i,t}$ is the difference of A-share close price and H-share close price over H-share close price, after adjusting the exchange rate on that trading day:

$$Premium_{i,t} = \frac{Close\ Price\ \ A_{i,t} - Close\ Price\ \ H_{i,t}}{Close\ Price\ \ H_{i,t}} \qquad (5.2)$$

The conversion of exchange rates is automatically calculated by the data source. It is the A-H premium of a certain company on a certain trading day. Out of concerns of omitted variables, we include the first lag term of premium $Premium_{i,t-1}$ as a control variable. It may have correlations with sentiment variables and impose omitted variables bias. For example, yesterdays' premia might affect investors beliefs on

today's market return. If yesterdays' premia is high, investors may be more optimistic. We use the trading volume of call options on SSE 50 Index as a proxy for optimistic beliefs, and the trading volume of put options on SSE 50 Index as a proxy for pessimistic beliefs. SSE 50 Index are a stock index constructed by the Shanghai Stock Exchange, which represents the fifty largest and most liquid stocks on the Shanghai Stock Exchange. The underlying assets can be deemed as a miniature of the A-share market, though it doesn't include stocks with small market capital, and it misses certain industries. The selection of this pair of options is reasonable, for they represent opposite beliefs on the market. There can be other proxies, but we believe that our selection is direct and clear-cut. $\varepsilon_{i,t}$ is the error term, which includes other factors that affect A-H premium but is uncorrelated with other dependent variables. For instance, it can contain different risk premia. The noise term is assumed to have an expectation of zero. The following models also use the same notations, including subscripts and variable names.

According to our theory, we would expect a significantly positive coefficient of optimistic beliefs and an insignificant coefficient of pessimistic beliefs in model (5.1), as we would expect the short-sell constraints mitigate the pessimistic beliefs and optimistic beliefs trigger noise traders to drive the price up. In our assumption, noise traders act on their misbeliefs. If noise traders have stronger optimistic beliefs than pessimistic beliefs, they will bid up the share prices. In addition, we would expect the coefficient of trading volume of put options on SSE 50 Index to be negative, as it represents pessimistic beliefs which are supposed to lower prices, though it shouldn't be statistically significant. We set our significance level at 5%. Mei et al (2005) demonstrate that limits to arbitrage stimulate speculative motive in the A-share market with A-B premia, so we won't repeat the same process.

Second, we want to examine that current level of A-H premia contains information about future volatility: when A-H premia are high, they suggest high volatility in the future. If A-H premia contain such information, we see that arbitrageurs can read the future market volatility from A-H premia. As they learn that the market will be volatile in the future, out of risk aversion, they will choose to be more conservative and give up some of the arbitrage opportunities. Hence, the inflated share prices can't be corrected immediately. Different sources can trigger future market volatility, including noise trader risk. Model (5.3) examines that A-H premia suggest high volatility in the future.

$$Volatility_{t+\tau} = Volatility_t + Premium_{i,t} + \varepsilon_{i,t} \quad (5.3)$$

We use implied volatility of SSE 50 Index call and put options as our market volatility variable. It's backed out with Newton-Raphson iteration method by the data provider, CSMAR. The method is based on Black-Scholes option pricing model. As there are different call and put option contracts on the market, we take the average of the implied volatility of different option contracts as our market volatility variable. We use the next τ period implied market volatility $Volatility_{t+\tau}$ to represent future market volatility. $Volatility_t$ is included for robustness reason similar to the one we have discussed above on $Premium_{i,t-1}$.

Based on our theory we would expect a significantly positive coefficient of A-H premium on future market volatility. When A-H premia are at a high level, they imply high market volatility in the future in A-share market. Arbitrageurs also perceive this risk.

To see how far the arbitrageurs can perceive, we run model (5.3) with τ from 1 to 1095 (three years) repeatedly and record the coefficients and t-statistics. If high A-H premia imply high market volatility to a long term, it strengthens our argument that arbitrageurs will fail to correct mispricing in the short run. When the expected future market volatility persists, arbitrageurs can't find intervals to jump in and correct the mispricing as soon as possible when mispricing appears.

Third, we want to demonstrate that optimistic noise trading stimulates future market volatility, a noise trader risk. We set up model (5.4) for examination:

$$Volatility_{t+\tau} = Volatility_t + Optimistic_{i,t} + \varepsilon_{i,t} \quad (5.4)$$

Noise trading causes future market volatility. When a surge of optimism appears in the market, it first amplifies itself through positive-feedback strategies and later gets

frustrated when noise traders tread on each other. In that case, we expect to see in model (5.4) that the coefficient of $Optimistic_{i,t}$ is positive and significant.

In order to see for how long optimistic noise trading can ruffle the market, we run model (5.4) repeatedly and record the coefficients of $Optimistic_{i,t}$ and the corresponding t-statistics, as how we iterate model (5.3) with different τ values. We expect a coefficient pattern similar to that of model (5.3), as noise trader risk is one of the sources of future market volatility, so noise trader risk should show similar pattern with expected market volatility. When noise trader risk is at a high level, the future market risk is supposed to be high as well.

Forth, we want to show noise trader risk inflates A-share prices to fortify our argument. We regress $Volatility_{t+1}$ on $Optimistic_{i,t}$ to predict the expected market volatility on the following day and test whether the predicted future market volatility have a positive impact on A-H premia.

$$E(Volatility_{t+1}) = Volatility_t + Optimistic_t \tag{5.5}$$

$$Premium_{i,t} = Premium_{i,t-1} + E(Volatility_{t+1}) + \varepsilon_{i,t} \tag{5.6}$$

We expect a positive and significant coefficient of expected market volatility on A-H premia in model (5.6).

These four models conclude our explanation for A-H premia with noise trader risk. Model (5.1) tests that optimistic beliefs is the main power driving noise trading, compared to pessimistic beliefs, which might be mitigated by limits of short-sales constraints. In model (5.3), we examine whether high A-H premia foretell high market volatility in the future, which overawes risk averse arbitrageurs. Model (5.3) and (5.4) demonstrate noise trader risk is a source of expected market volatility. Model (5.4) and (5.6) examine whether there is noise trader risk and whether it inflates A-share prices.

We use an unbalanced daily data sample in the period from November 2014 to November 2019. The period is carefully chosen with consideration of impact of policy shift and a special event. Our data sample begins on 23rd November 2014, right after the launch of Shanghai Hong-Kong Stock Connect program, which is supposed to a

mechanism that allows investors on both markets, mainland China and Hong Kong, to trade eligible shares on the other market via brokers. This policy is supposed to reduce A-H premia. To avoid the disruption of this event, we choose to a data sample beginning after the event. Though we process the data up till February 2020, we decide to drop the portion after 23rd November 2019, a month before an influential event. On 23rd December 2019, when China's stock market further relieved short sell constrains when the China Financial Futures Exchange released options on China Stock Index 300, the main stock index of China's stock market. This event is supposed to relieve limits to arbitrage and disrupt our empirical research. We deem the level short sell constrains as invariant, so we need to drop the later data sample. Our data sample includes seventy-three A-H dual-listing companies in total with 73576 observations over time.

Our data sources are the CSMAR database and Choice Financial Terminal. We retrieve the option data from CSMAR and the A-H premia data from the Choice Financial Terminal. Both sources are credit worthy. CSMAR is a database focus of China's financial and economic data. Data from CSMAR has been used in top journals in finance, such as Journal of Finance. CSMAR has also been directed by the Wharton Research Data Services as the sole data source. Choice Financial Terminal is a terminal used by domestic institutions for professional services. Choice Financial Terminal's data is from different stock exchanges directly, so shifting to other finical data providers won't bring data quality improvement.

The models are estimated with panel OLS, with control of individual effects of different dual-listing companies. Time effects shouldn't be controlled, as our main independent variables are time effects. To address robustness concerns, we include aforementioned control variables to reduce the impact of omitted variable bias. We adopt a 5% significance level for all our examination.

To demonstrate the effect of limits to arbitrage, we examine a special event on 23rd December 2019, which was supposed to relieve limits to arbitrage. As mentioned above, on 23rd December 2019, the China Financial Futures Exchange launched option products on China Stock Index 300. China Stock Index 300 is one of the most representative indexes in the A-share market, which include the three hundred largest stocks in the market. Qualified investors can short sell CSI 300 index with the new product, even though the qualification bar most of individual investors from it.

According to the China Financial Futures Exchange, investors with more than a minimum available balance of RMB 500,000 in its margin account for more than 5 consecutive trading days can trade options. This event relieves limits to arbitrage to arbitrageurs.

We include a dummy variable $Event_t$ to represent this event and interact $Optimistic_t$ with $Event_t$ to see if reliving limits to arbitrage in model (5.7) dampens the impact of optimistic beliefs.

$$Premium_{i,t} = Premium_{i,t-1} + Optimistic_t + Pessimistic_t$$
$$+Event_t + Optimistic_t \times Event_t + \varepsilon_{i,t} \qquad (5.7)$$

In this regression, we expect the interaction term $Optimistic_t \times Event_t$ to be significantly negative. According to our theory, limits to arbitrage mitigate pessimistic beliefs and therefore optimistic beliefs are more pronounced. When short-sales constraints are relieved, pessimistic beliefs will be less oppressed and optimistic beliefs becomes less pronounced relatively.

To avoid other interrupting events and keep a decent number of observations, we use part of the data sample mentioned above. We use the data from February 2019 to February 2020, a one-year long data set. Except for the time frame, this data set is the same as the previous dataset we use in other aspects. This examination is only for qualitative demonstration and it suffers from several drawbacks. For example, other unobserved events may have an impact on A-H-share premia and on market sentiments. In addition, it might take a while for the new option product to kick off, or the market may have already perceived this event way before the launch of the product. Considering these potential defects, we only include this regression for demonstration purpose in support of our argument. Previous papers, such as Mei et al's (2005) paper, demonstrate this argument well enough in A-B premia. We include this event analysis only to supplement our research.

6. *Empirical Results*

In this section, we examine our regression results. We include our results in the following tables and figures and show the empirical results support our noise trader risk theory in explaining A-H premia.

In model (5.1), we have a significantly positive coefficient of optimistic beliefs and an insignificant coefficient of pessimistic beliefs. It demonstrates that optimistic beliefs are more influential in the A-share market and the optimistic variable has a positive impact on A-H premia. The signs of coefficients defend our selection of investor sentiment proxies as well. The positive sign of the optimistic belief coefficient shows that when the trading volume of the call option is high, the A-share prices are further inflated. In comparison, the negative sign of the pessimistic belief coefficient shows that pessimistic beliefs narrow the A-H price gap, though it is not significantly different from zero. In terms of mitigation of pessimistic sentiment, Mei et al (2005) demonstrate that limits to arbitrage stimulate speculative motive in their A-B premia research. Hence, the regression results of model (5.1) indicates that limits to arbitrage pronounce optimistic beliefs, while oppress pessimistic beliefs in the market. It also demonstrates that optimistic noise trading inflates A-share prices.

In model (5.3), we see that A-H premia foretell future market volatility significantly. When A-H premia are high, they foretell high market volatility in the future. From the perspective of arbitrageurs, when observing high A-H premia, they learn at the same time that there will be high volatility in consecutive trading days. To them, the swings are a sword with two blades, as the price may rise even further against their interests. Therefore, out of risk aversion, arbitrageurs will become more conservative when the prices are highly inflated.

Figure 6.1 shows how the expected high volatility persists over a long period of time. The red line is the coefficient of A-H premium and the blue line is its t-statistics. As the horizon goes forward, the coefficient of A-H premium becomes larger and more significant, reaches its maximum in the middle term (around one year), and sharply drops to insignificant in the later period. It demonstrates that the current inflated price level contains information about a long-lasting market volatility in the future. When arbitrageurs expect a volatile period lasting a long time, they won't step in and correct

the mispricing aggressively until the market become quite again, for they dislike taking risk. Figure 6.1 suggests that arbitrageurs are unlikely to fully correct mispricing immediately, as the high market volatility is expected to last for more than one year. It supports our argument that arbitrageurs are overawed by perceived future volatility and fail to correct mispricing.

Empirical results of model (5.4) show that optimistic beliefs have a positive and significant impact on market volatility in the future. In our assumption, noise traders act on noise, or their misbeliefs, and trade. They will amplify the market volatility through either positive-feedback trading or through treading on each other to escape. Therefore, the statistical results reveal that noise traders stimulate market volatility in the future, the creation of noise trader risk.

Figure 6.2 shows how long noise trading can stimulate future market volatility in the market. The red line is the coefficient of A-H premium and the blue line is its t-statistics. As the horizon goes further into the future, the coefficient of optimistic beliefs on future market volatility gets higher and reached its maximum in the middle term (around one year). It shows that noise trading can create high volatility in the market for a long period of time. An interesting point in Figure 6.2 is that on around day 800, the coefficient of optimistic beliefs turns negative. It suggests that, in the long term, the impact of optimistic beliefs fades away and so does noise trader risk. It's reasonable and conforms to reality. When a surge of optimism appears in the market, it first amplifies itself through positive-feedback strategies and later gets frustrated when the price skyrockets too high and noise traders tread on each other. Either way, the market volatility is amplified. When the enthusiasm and the aftermath are gone, the market becomes quite again.

Figure 6.2 show a coefficient pattern similar to Figure 6.1. It suggests that noise trader risk is a source of expected market risk in support of our argument. The trajectory of optimistic belief coefficient and that of premium coefficient increase in the beginning and reach their maximum in the middle term (around a year). However, when the horizon goes beyond one-year time, the two figures show different pattern. The coefficient of A-H premium begins to decrease, while the coefficient of optimistic beliefs still persists on the plateau and suddenly plummets towards year three. It may suggest, in a longer horizon, noise trader risk contributes less to the expected market

volatility, and investors value other sources of volatility more when they expect the three year ahead volatility. It makes sense in that investor sentiment is more short-lived than other market influencing factors. Combining the two figures and the empirical results of model (5.4), we can see that noise trader risk stimulates expected market volatility for a fairly long period of time.

In model (5.6), we see a positive and significant coefficient of expected market volatility on A-H premia. It suggests that noise trader risk inflates A-share prices. It's reasonable argue that arbitrageurs expect the noise trader risk and miss arbitrage opportunities out of risk aversion. Due to the lack of price correction, the inflated prices keep hovering over the fundamental prices. That forms A-H premia. This result completes our noise trader risk theory.

Putting all the empirical results together, we have evidence to support our argument that, in the presence of limits to arbitrage, noise trader risk inflates A-share prices and creates A-H premia. The high R-square values in the regression results are reasonable, as the control variables have Markov property in a daily data sample. In model (5.1), the current level of A-H premia has a coefficient of 0.99; in model (5.4), the current level of market volatility has a coefficient of 0.95. This shows that today's level depends highly on last day's level. In addition, this suggests that these control variables are necessary to capture omitted variable bias.

Table 6.1

$$\boldsymbol{Premium_{i,t} = Premium_{i,t-1} + Optimistic_t + Pessimistic_t + \varepsilon_{i,t}} \quad \textbf{(5.1)}$$

The dependent variable $Premium_{i,t}$ is the difference of A-share close price and H-share close price over H-share close price, after adjusting the exchange rate on that trading day, defined as $Premium_{i,t} = Close\ Price\ \ A_{i,t} - Close\ Price\ \ H_{i,t}\ /Close\ Price\ \ H_{i,t}$. Optimistic belief $Optimistic_t$ is measured by trading volume of call options on SSE 50 index, and pessimistic belief $Pessimistic_t$ is measured by trading volume of put options on SSE 50 index. $Premium_{i,t-1}$is included for the robustness reason. Subscript i stands for individual stocks, and subscript t stands for time. Daily CSMAR data from 2014 to 2019.

Variable	Coeff.	Std.err	T-stat	p-value	R_2
LagPremium	0.9974	0.0002	4021.7	0.0000***	
Optimistic	3.56e-09	1.29e-09	2.7607	0.0058***	
Pessimistic	-1.992e-09	1.568e-09	-1.2707	0.2038	
					0.9967

Table 6.2 premium coefficient

$$\boldsymbol{Volatility_{t+\tau} = Volatility_t + Premium_{i,t} + \varepsilon_{i,t}} \quad (5.3)$$

This model is to examine our hypothesis based on above four models. This table shows the premium coefficients.

Horizon	Coeff.	std.err	T-stat	p-value	R_2
1 day	0.0343	0.0010	33.652	0.0000***	0.9435
1 week	0.0995	0.0017	59.342	0.0000***	0.8462
1 month	0.1113	0.0018	62.308	0.0000***	0.8229
1 year	0.1337	0.0026	52.346	0.0000***	0.6665
2 years	0.0615	0.0022	27.589	0.0000***	0.8381

Table 6.3

$$\boldsymbol{Volatility_{t+\tau} = Volatility_t + Optimistic_{i,t} + \varepsilon_{i,t}} \quad (5.4)$$

The dependent variable $Volatility_{t+\tau}$ is next period implied market volatility, measured by implied volatility of SSE 50 index call and put options. $Volatility_t$ is included by for robustness reason. Optimistic belief $Optimistic_t$ is measured by trading volume of call options on SSE 50 index. Subscript i stands for individual stocks, and subscript t stands for time. Daily CSMAR data from 2014 to 2019.

Variable	Coeff.	Std.err	T-stat	p-value	R_2
Volatility	0.9548	0.0011	870.69	0.0000***	
Optimistic	2.26e-08	8.701e-10	25.971	0.0000***	
					0.9431

Table 6.4

$$\boldsymbol{Premium_{i,t} = Premium_{i,t-1} + E(Volatility_{t+1}) + \varepsilon_{i,t}} \quad (5.6)$$

The dependent variable $Premium_{i,t}$ is the difference of A-share close price and H-share close price over H-share close price, after adjusting the exchange rate on that trading day. $E(Volatility_{t+1})$ is defined as $E(Volatility_{t+1}) = Volatility_t + Optimistic_t$. Subscript i stands for individual stocks, and subscript t stands for time. Daily CSMAR data from 2014 to 2019.

Variable	Coeff.	Std.err	T-stat	p-value	R_2
LagPremium	0.9971	0.0003	3182.7	0.0000***	
Predicted Volatility	0.0024	0.0004	5.5497	0.0000***	
					0.9966

Table 6.5

$$\boldsymbol{Premium_{i,t} = Premium_{i,t-1} + Optimistic_t + Pessimistic_t + Event_t + Optimistic_t \times Event_t + \varepsilon_{i,t}} \quad (5.7)$$

The dependent variable $Premium_{i,t}$ is the difference of A-share close price and H-share close price over H-share close price, after adjusting the exchange rate on that trading day, defined as: $Premium_{i,t} = (Close\ Price\ \ A_{i,t} - Close\ Price\ \ H_{i,t})/Close\ Price\ \ H_{i,t}$. Optimistic belief $Optimistic_t$ is measured by trading volume of call options on SSE 50 index, and pessimistic belief $Pessimistic_t$ is measured by trading volume of put options on SSE 50 index. $Event_t$ is a dummy variable to represent the event in model (5.1). Subscript i stands for individual stocks, and subscript t stands for time. Daily CSMAR data from 2014 to 2019.

Variable	Coeff.	Std.err	T-stat	p-value	R^2
LagPremium	0.9993	0.0005	2217.8	0.0000***	
Optimistic	-5.268e-10	1.102e-09	-0.4778	0.6328	
Event	0.0020	0.0016	1.2542	0.2098	
Interaction	-2.054e-09	1.051e-09	-1.9542	0.0507*	
Pessimistic	2.192e-09	1.336e-09	1.6410	0.1008	
					0.9982

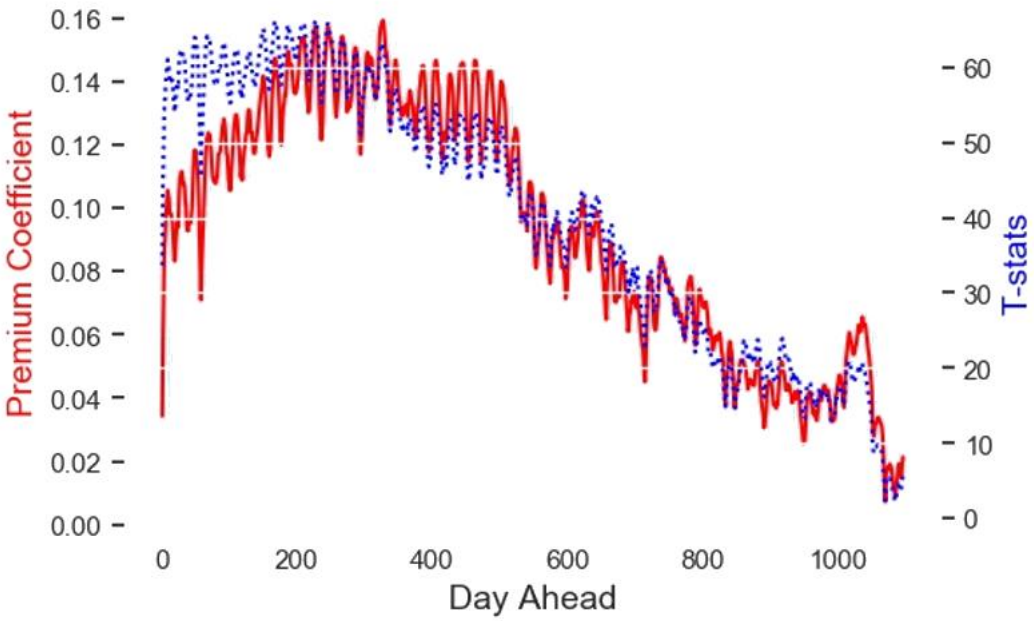

Figure 6.1

The figure shows the expected high volatility persists over a long period of time. The red line is the coefficient of A-H premium and the blue line is its t-statistics.

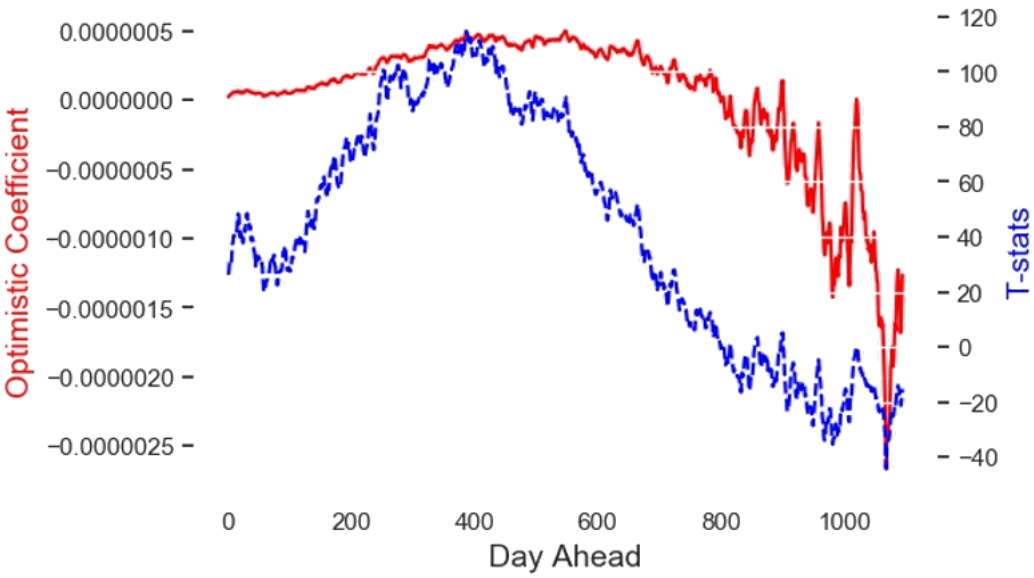

Figure 6.2

The figure shows how long noise trading can stimulate future market volatility in the market. The red line is the coefficient of A-H premium and the blue line is its t-statistics.

In our supplementary examination of the special event, we find that the interaction term between optimistic beliefs and the event has a negative and significant coefficient, which is in conformity with our expectation. This may suggest relief of limits to arbitrage dampen the effect of optimistic noise trading. With the data sample shortened, we see that the coefficients of the event, optimistic beliefs are not significant, but they are significant in the previous dataset that covers a longer period. The insignificant coefficient of the event may suggest this event is not impactful on A-H-share price discrepancy. The empirical results of model (5) are vulnerable to robustness check. They are provided only for demonstration, as we mention before. It can be improved with other proxies for the strongness of short sales constraints, for example, breadth of companies or the TED in China's stock market.

7. Competing Hypotheses in the Literature

In this section, we summarize four competing hypotheses in the literature. The relation between our paper and these literatures is discussed in Section 2.

7.1 Differential Risk Hypothesis

Differential risk hypothesis is clear-cut and easy to understand from the efficient market point of view. Different groups of investors require different compensations for risk-taking in separated markets, due to different risk aversion. China's stock market is immature compared to well-developed stock markets like the US's stock market, and domestic investors are isolated from the other stock markets because of strict government restrictions. This can lead to different perceptions of risk and thus risk aversion. Ma (1996) examines that A-B-share stock price discrepancy stems from distinctive investors' risk aversion whereas B-share and foreign share discrepancies are from differentiated expectations. Sun and Tong (1999) suggest that domestic investors in China are more optimistic than foreign investors on expected growth and their optimistic attitude is a reason for A-B-share premia. Li et al (2006) find that the A-H premia are associative with the contemporaneous movements of H-share market index relative to A-share market index with a data sample form January 1997 to March 2002. Another way to perceive differential risk premia is that there might be special risks involved in investments in China's stock market, such as risk of company expropriation by the government (Guo et al, 2013), political risk (Karolyi et al, 2009), and so on. There might be other similar factors, controlling for discovered risk premia.

7.2 Differential Demand Hypothesis

Differential demand hypothesis implies that different groups of investors have different demand for the class of shares they are allowed to purchase, and thus the prices are different. In their theoretical model, Stulz and Wasserfallen (1995) demonstrate that the different prices between foreign shares and domestic shares are the results of price discrimination, which takes advantage of binding ownership restrictions and can maximize the firm value. In their empirical evidence, foreign investors need to pay

more than domestic investors for Swiss companies' stock. It's quite the opposite to A-H-share premia. In China's stock market, domestic investors face fewer investment options than foreign investors, so domestic investors may have lower required rates of turn (Fernald and Rogers 2002). They examine both A-H-share and A-B-share premia and suggest that the premia and high volatility can only be explained with the CAPM model, if domestic investors require lower expected returns.

7.3 Liquidity Hypothesis

Liquidity hypothesis argues that the stock price disparity is due to the difference of liquidity of the same assets in segmented market. It suggests that premia can be explained by a compensation for illiquidity. Bailey (1994) and Poon et al (1998) suggest the discounts of B-shares to A-shares are caused by the illiquidity of B-shares. The B-share market is known for its illiquidity, therefore it's reasonable to suggest that discounts are given to compensate for such a risk to attract foreign investors. In addition, trading costs in the H-share market are generally higher than in the A-share market, which presumably will make the H-share less liquid. Lee (2009) revisits the liquidity hypothesis with regard to A-H premia and finds evidence suggesting that A-shares are in general more liquid than their parallel H-shares and liquidity can be used to explain for A-H premia, as the spread and depth can significantly explain for the disparity, controlling for other risk premia.

7.4 Asymmetric Information Hypothesis

Asymmetric information hypothesis suggests that the price differences are caused by asymmetry information received by different groups of investors. One of the groups might be better informed than the other with respect to the underlying company, due to langue advantages, better understanding of local accounting standards, etc. Lack of information may require a discount. Easley et al (2002) suggest that information affects asset prices though showing that the probability of information-based trading affects stock expected returns: the less informed the investors are, the less they are willing to pay for an asset. Li and Lu (2014) show the similar results in China's A-share market. Chakravarty et al (1998) suggest that foreign investors have less information because

of insider trading and share manipulation, and they empirically examine their argument with A-B premia in the period from in November 1990 to July 1991. On the contrary, in their later research, Chan et al (2008) indicates that domestic investors are in an information disadvantage and such asymmetry can explain for A-H premia: when domestic investors are allowed to trade B-shares, the information disadvantage decreases and so do the B-share discounts. Yang (2003) conducts recursive cointegration analysis and indicates that foreign investors better information access towards emerging markets over domestic investors thus leading to A-B-share discrepancies. These explanations are contradictory in that acquiring more information can lead to either a higher price that accounts for confirmed great prospects of the company or a lower price for fully understanding of the underlying risk.

8. Discussion and Policy Implication

In this section, we demonstrate our empirical results of the alternative hypothesis and discuss about policy implication based on those results and former sections. In the end, we combine our findings with investment suggestions.

To complete our discussion about A-H premia, we re-examine previous hypotheses for A-H premia puzzle. The competing hypothesis are inconsistent with our data sample. Referring to previous literature, we replicate the model of Chan et al (2001) with modified variables and an updated data sample.

$$Premium_{i,t} = Premium_{i,t-1} + Turnover\ A_{i,t} + Turnover\ H_{i,t} + State\ Ownership_{i,t} + AH\ Outstanding\ Ratio_{i,t} + Asset\ Float_{i,t} + AH\ Volitility\ Ratio_{i,t} + A\ Market\ Index_t + H\ Market\ Index_t + \varepsilon_{i,t} \quad (8.1)$$

The variables are introduced below, and the description of the data is in the appendix:

Variable	*Description*
$Premium_{i,t}$	A-H premium
$Turnover\ A_{i,t}$	Daily turnover of A-share
$Turnover\ H_{i,t}$	Daily turnover of H-share
$State\ Ownership_{i,t}$	Percentage of state ownership on a board basis
$AH\ Outstanding\ Ratio_{i,t}$	Number of outstanding H-shares over A-shares
$Asset\ Float_{i,t}$	Total Value of outstanding A-shares and H-shares
$AH\ Volitility\ Ratio_{i,t}$	Twenty-day volatility of A-share over H-share
$A\ Market\ Index_t$	A-share market index - CSI 300
$H\ Market\ Index_t$	H-share market index - HSI

Table 8.1

$$\boldsymbol{Premium_{i,t} = Premium_{i,t-1} + Turnover\ A_{i,t} + Turnover\ H_{i,t} + State\ Ownership_{i,t} + AH\ Outstanding\ Ratio_{i,t} + Asset\ Float_{i,t} + AH\ Volitility\ Ratio_{i,t} + A\ Market\ Index_t + H\ Market\ Index_t + \varepsilon_{i,t}} \quad \textit{(8.1)}$$

The dependent variable $Premium_{i,t}$ is the difference of A-share close price and H-share close price over H-share close price, same as above. $Turnover\ A_{i,t}$ and $Turnover\ H_{i,t}$ are pairwise liquidity of A-shares and H-shares.$State\ Ownership_{i,t}$ is proxy for political risk. $AH\ Outstanding\ Ratio_{i,t}$ represents differential demand. $Asset\ Float_{i,t}$ shows asymmetric information hypothesis and $AH\ Volitility\ Ratio_{i,t}$ relative risk tolerance. $A\ Market\ Index_t$ and $H\ Market\ Index_t$ represent differential risk hypothesis. Subscript i stands for individual stocks, and subscript t stands for time. Daily CSMAR and choice data from 2014 to 2019.

Variable	Coeff.	Std.err	T-stat	p-value	R^2
LagPremium	0.9990	0.0009	1082.6	0.0000***	
TurnoverA	0.1708	0.0283	6.0294	0.0000***	
TurnoverH	0.1876	0.0559	3.3561	0.0008***	
State Ownership	-0.0129	0.0041	-3.1266	0.0018***	
AH Outstanding Ratio	-0.0049	0.0027	-1.7854	0.0742*	
Asset Float	8.068e-10	1.281e-09	0.6299	0.5288	
AH Volatility Ratio	0.0013	0.0010	1.3417	0.1798	
A Marekt Index	0.0081	0.0004	20.917	0.0000***	
H Market Index	-0.0092	0.0005	-18.747	0.0000***	
					0.9973

These variables are proxies for different theories that we discuss in the literature review section. $Turnover\ A_{i,t}$ and $Turnover\ H_{i,t}$ represent the pairwise liquidity of A-shares and H-shares, proxies for liquidity hypothesis. The higher the turnover rate is, the better the liquidity. To examine political risk, we use $State\ Ownership_{i,t}$ as a proxy. If the share holdings are concentrated to the state, we see there is more involved political risk. $AH\ Outstanding\ Ratio_{i,t}$ is used as a proxy for differential demand hypothesis. The higher this ratio is, the lower the foreign demand is or the higher the domestic demand is. It can only demonstrate whether this hypothesis hold. We use $Asset\ Float_{i,t}$ to account for the asymmetric information hypothesis, with the assumption that if a company is larger in scale, its information is better received by

investors. To capture the relative risk tolerance in the two markets, we include $AH\ Volitility\ Ratio_{i,t}$. Market indexes, $A\ Market\ Index_t$ and $H\ Market\ Index_t$, are used to examine differential risk hypothesis.

We run the regression with the data sample from 2014 to 2019, which is the same period we use in the last section. The data is also from CSMAR and Choice Financial Terminal. The derivation and calculation of the data are described in the Appendix. We use panel OLS to estimate the parameters in conformity with the last section.

The empirical evidence suggests that only differential risk hypothesis holds, as the coefficient of A-share market index is significantly positive, and the coefficient of H-share market index is significantly negative. For the liquidity hypothesis, when liquidity in H-shares is high, the A-H premia should be lower, as the required discounts in H-shares are lower, but the result suggests against it. The coefficient of H-share turnover rate is positive. For political risk, the coefficient of state ownership is not positive, the opposite to our expectation. We expect the higher political risk is, the larger A-H premia are. For the differential demand hypothesis, we expect the coefficient of H-share relative demand to be significantly negative, but it's not significant at the 5% significance level. For the asymmetric information hypothesis, the coefficient of asset float is not significant. In our expectation, it should be significantly negative to support this hypothesis, for if the information is better received by different investors, the price discrepancy between A-share and H-share is supposed to be narrower. In terms of risk aversion, we expect the relative A-share volatility to H-share volatility to have a significantly positive coefficient, as we expect that if domestic investors in the A-share market are more risk tolerant than investors in the H-share market, increase of relative A-share volatility will increase the price disparity. Of course, these variables may not be the best to represent the hypothesis, and a new set of variables may lead to different results. It remains to be further examined with other variables.

In the literature review view, we see that Fan and Wang (2016) find empirical evidence in support of the differential risk hypothesis and liquidity hypothesis. Our results are different from theirs. One reason is that their data sample includes the period when Shanghai Hong-Kong Stock Connect program was launched. The other reason is that we use a different set of independent variables. Despite of the different results, the

evidence suggests that the differential risk hypothesis holds when it comes to explaining for A-H premia. Now we learn that A-H premia can be explained by differential risk hypothesis and noise trading, but can we draw any policy implications?

The first question to answer is: should the related authorities try to converge A-share and H-share prices? According to our findings, A-H premia can be explained by different risk attitudes and by optimistic noise trading distortions. Different risk attitudes harbored by different groups of investors may reflect the lack of transparency and accessibility to company and market information in the A-H-share market. Indeed, listed A-share companies only reveal their financial intuitions in Chinese financial reports, which sometimes are not reliable with suspicious book-cooking. When financial reports barely reveal the operation situations of companies, investors can hardly see the fundamental value within. Shrewd investors will choose to leave the market and leave behind speculators and noise traders. Speculative atmosphere is fermented. Mispricing derived from noise trading is not a good sign to the A-share market. It distorts the allocation of capital in the A-share market and leads it to market speculation. In chasing surges of optimism, capital is attrite by transaction costs, instead of flowing into companies in the form of machinery, lands, or labors. With these two problems, the pricing and capital allocation functions in the A-share market would be impaired. In the long term, investors would either leave the market complete to other markets and fund-seeking companies would turn to borrowing. Overvaluation in the A-share market is an issue to be solved.

The next question is how to solve this mispricing problem. Our findings suggest that limits to arbitrage mitigate pessimistic beliefs and pronounce optimistic sentiment. It's tempting to recommend the A-share market relieve limits to arbitrage all together at once, but the A-share market is vulnerable to a dramatic change as its legal and finical system is waiting to be further improved. As most of the market participants are individual investors, it's advisable to educate inexperience investors. Basic investment knowledge should be imparted thoroughly to investors, instead of superficially, when they open a security account. In addition, the transparency of listed companies needs to be improved. Companies should be encouraged or stipulated to reveal their financial reports in other languages, for foreign investors to access and supervise. Meanwhile, further efforts should be made to introduce foreign investors into the A-share market, who are supposed to be more experienced than domestic investors in terms of exposure

to economic cycles more developed markets. Big money should guide small money, as foreign and domestic institutions guide individual investors. Legal and financial system should be refined to crack down market manipulation and insider trading.

In the long run, the design of the A-share market needs to be reconsidered. First, the standpoint should change from funding state-run companies to funding good companies. The procedure of granting IPOs should be simplified and reduced its bias towards state-run companies. When it is appropriate, the procedure should shift from approval-based IPO system to registration-based IPO system, but the concern here is that individual investors are inexperienced and can't access as much information on the company waiting for IPO as the committee that gets to grant an IPO. Investors education is much needed. Second, the market exit system in the A-share market should be completed. In fact, capital is attracted towards suspended stocks, which have a layer of shell value. These suspended companies with shell value are waiting to be salvaged by companies which want to get listed but don't bother to go through the cumbersome IPO procedure. Investors pour in their money, expecting that the shell value will be realized sometime. The consensus is that stocks in the A-share market will never be forced to exit. A complete exit system will inform investors that such speculative strategies might fail. Regarding media control, should the related authorities loosen the control of views? We learn that government censorship which holds negative news in control helps to inflate A-share price prices (Dong et al, 2018), but we don't know whether removing it will cause dissemination of market manipulation rumors. It's safer to suggest keeping the status quo.

From the perspective of investors, our findings are meaningful. We find that A-H premia is a good indicator of market volatility in the future. When the market spirit is high, there will be large volatility in the future persisting a fairly long period of time, roughly one year. If market conditions allow, a straddle strategy betting against the volatility in the A-share market is recommended. In practice, if a positive shock hits the market and is confirmed by A-H premia, a speculator can form a straddle portfolio to bet against the A-share market volatility. Such a portfolio can be constructed with SSE 50 and CSI 300 call and put options.

9. *Conclusion*

A-H premia is an intriguing puzzle. It is large in magnitude and persistent, despite related authorities' market de-segmentation attempts. Previous literature based on various forms of market segmentation hypotheses cannot successfully explain for the great magnitude of A-H premia. Furthermore, we find that those explanations are inconsistent with our data in the sample.

We adopt a simple theoretical explanation based on noise trader risk and limits to arbitrage to explain for A-H premia, and empirically examine this hypothesis with A-H premia data sample from 2014 to 2019. We argue that arbitrageurs avoid correcting mispricing because of expected future price volatilities driven by investor sentiment. First, we find that A-H premia is significantly correlated with the measure of optimistic investor sentiment. In addition, we examine the existing hypotheses in the literature and find that only differential risk hypothesis is supported by our data sample and only 57% of the variation of A-H premia can be explained by the various previous hypotheses. Second, we find that A-H premia has a significant predictive power of future stock market volatility. Interestingly, the predictability lasts for over one year. This indicates that arbitrageurs avoid correcting mispricing because of expected future price volatilities in one-year horizon. Third, we find that the correlation between A-H premia and investor sentiment becomes stronger during the period with higher limits to arbitrage.

Inflated A-share prices are problematic, for they distort the asset allocation strategy and effective pricing functions of the A-share market. Reformation is needed to address this problem. We recommend improving investor education, opening the market further to foreign investors, and strengthening the listing and delisting system. From the practical point of view, we find that A-H premia level is a good indicator of market volatility in the future.

References

Ali F.Darrat, Maosen Z., 2010, On the Chinese B-share price discount puzzle: some new evidence, Journal of Business Research 63(8), 895-902.

Amihud Y., Mendelson H., 1986, Liquidity and stock returns, Financial Analysts Journal 42(3), 43-48.

Bailey W., 1994, Risk and return on China's new stock markets: Some preliminary evidence, Pacific-Basin Finance Journal 2(2), 243–260.

Black F., Noise, The Journal of Finance, 1986, 41(3), 528-543.

Biqiong Zhang, Li Yue, 2002, Do the exchange rates matter for Chinese stock market: evidence from ARDL-ecm estimation, Journal of Finance.

Burdekin R C., Redfern L., 2009, Stock market sentiment and the draining of China's savings deposits, Economics letters 102(1), 27-29.

Cai C X., McGuinness P B., Zhang Q., 2011, The pricing dynamics of cross-listed securities: the case of Chinese A-and H-shares, Journal of Banking & Finance 35(8), 2123-2136.

Chakravarty S., Sarkar A., Wu L., 1998, Information asymmetry, market segmentation and the pricing of cross-listed shares: Theory and evidence from Chinese A and B-shares, Journal of International Financial Markets, Institutions and Money 8(3-4), 325-356.

Chan K., Kwok J K., 2005, Market segmentation and share price premium: evidence from Chinese stock markets, Journal of Emerging Market Finance 4(1), 43-61.

Carpenter, Jennifer N., and Robert F. Whitelaw. "The development of China's stock market and stakes for the global economy." Annual Review of Financial Economics 9 (2017): 233-257.

Chen J., Hong H., Stein J C., 2002, Breadth of ownership and stock returns, Journal of Financial Economics 66(2-3), 171-205.

Chen G., Lee B., Rui O., 2001, Foreign ownership restrictions and market segmentation in China's stock market, Journal of Financial Research 24(1), 133–155.

Chen H., Chong T T L., She Y., 2014, A principal component approach to measuring investor sentiment in China, Quantitative Finance 14(4), 573-579.

Chung T., Li K., Cho-Hoi H., 2011, Explaining share price disparity with parameter uncertainty: Evidence from Chinese A- and H-shares, St. Louis: Federal Reserve Bank of St Louis.

Chung T K., Hui C H., Li K F., 2013, Explaining share price disparity with parameter uncertainty: evidence from Chinese A-and H-shares, Journal of Banking & Finance 37(3), 1073-1083.

De Long J B., Shleifer A., Summers L H., Waldmann R J., 1990, Positive feedback investment strategies and destabilizing rational speculation, The Journal of Finance 45(2), 379-395.

De Long J B., Shleifer A., Summers L H., Waldmann R J., 2005, Noise trader risk in financial markets, Journal of Political Economy 98(4), 703-738.

Eun C S., Janakiramanan S., 1986, A model of international asset pricing with a constraint on the foreign equity ownership, The Journal of Finance 41(4), 897-914.

Easley D., Hvidkjaer S., Ohara M., 2002, Is information risk a determinant of asset returns?, The Journal of Finance 57(5), 2185-2221.

Fernald J., Rogers, J H., 2002, Puzzles in the Chinese stock market, Review of Economics and Statistics 84(3), 416-432.

Fan Q., Wang T., 2017, The impact of Shanghai–Hong Kong stock connect policy on A-H-share price premium, Finance Research Letters 21, 222–227.

Fan H.K., Chia L.C., 2012, The determinants of overseas listing decisions: Evidence from Chinese H-share companies, Asian Business & Management 11, 591-613.

Gu M., Kang W., Xu B., 2018, Limits to arbitrage and idiosyncratic volatility: Evidence from China stock market, Journal of Banking & Finance 86, 240-258.

Guo L., Tang L., Yang S X., 2013, Corporate governance and market segmentation: evidence from the price difference between Chinese A and H-shares, Review of Quantitative Finance and Accounting 41(2), 385-416.

G M Chen, Bong-Soo Lee, Oliver Rui, 2014, Foreign ownership restrictions and market

segmentation in China's stock markets, Journal of Financial Research.

Hou W., Lee E, 2014, Split share structure reform, corporate governance, and the foreign share discount puzzle in China, The European Journal of Finance 20(7-9), 703–727.

Han X., Li Y., 2017, Can investor sentiment be a momentum time-series predictor? evidence from China, Journal of Empirical Finance 42, 212-239.

John A.D., Liu W.b., 2013, Information asymmetry, price discovery, and the Chinese B-share discount puzzle, Pacific-Basin Finance Journal 21(1), 1116-1135.

Jinghui Liu, Lan Alexander Eddie, 2007, Determinants of disclosure of A-, B- and H-share companies, Asian Reveiw of Accounting 15(1), 72-92.

Kalok Chan, Albert J., Menkveld, Zhishu Yang, 2008, Information asymmetry and asset prices: evidence from the China foreign share discount, Journal of Finance 63.

Karolyi G A., Li L., Liao R., 2009, A (partial) resolution of the Chinese discount puzzle, Journal of Financial Economic Policy.

Kumar, Anjali., Jun, Kwang, Saunders, Anthony, Selwyn, Susan, Sun, Yan, Vittas, Dimitri, Wilton, David, 1997, China's emerging capital markets, World Bank, 149-153.

Lakonishok J., Shleifer A., Thaler R., Vishny R., 1991, Window dressing by pension fund managers, National Bureau of Economic Research.

Lan Domowitz, Jack Glen, Ananth Madhavan, 2012, Market segmentation and stock prices: Evidence from an emerging market, The Journal of Finance, 1059-1085.

Lee H W., 2009, The price premium of China A-shares over Hong Kong H-shares: A further visit of the liquidity hypothesis, Asia-Pacific Journal of Financial Studies 38(5), 657-694.

Li G., Lu L., Wu B., Zhang Z., 2014, Asymmetric information, illiquidity and asset returns: evidence from China. Quantitative Finance 14(6), 947-957.

Li Y., Yan D., Greco J., 2006, Market segmentation and price differentials between A-shares and H-shares in the Chinese stock markets, Journal of Multinational Financial Management 16(3), 232-248.

Lior Menzly, Oguzhan Ozbas, 2010, Market segmentation and cross-predictability of returns, The Journal of Finance, 1555-1580.

Lin Tan, Thomas C., Chiang, Joseph R., Mason, Edward Nelling, 2008, Herding behavior in Chinese stock markets: An examination of A and B-shares, Pacific-Basin Finance Journal, 61-77.

Ma, 1996, Capital controls, market segmentation and stock prices, evidence from Chinese market, Pacific Basin finance Journal, 219-239.

Miller E M., Risk, 1977, Uncertainty and divergence of opinion, The Journal of Finance 32(4), 1151-1168.

Mei J., Scheinkman J., Xiong W., 2005, Speculative trading and stock prices: Evidence from Chinese AB-share premia, National Bureau of Economic Research.

Malkiel B G., Fama E F., 1970, Efficient capital markets: A review of theory and empirical work, The Journal of Finance 25(2), 383-417.

Mustafa N., Gultekin N., Bulent Gultekin, Alessandro Penati., 1989, Capital controls and international capital market segmentation: the evidence from the Japanese and American stock markets, Journal of Finance, 849-869.

Mei J., Scheinkman J., Xiong W., 2005, Speculative trading and stock prices: evidence from Chinese AB-share premia (No. w11362), National Bureau of Economic Research.

Pan J., Chi J., 2019, How does the Shanghai-Hong Kong stock connect policy impact the AH-share premium? Emerging Markets Finance and Trade, 1-17.

Poon W P., Firth M., Fung H G., 1998, Asset pricing in segmented capital markets: Preliminary evidence from China-domiciled companies, Pacific-Basin Finance Journal 6(3-4), 307-319.

Pekkat, Hietala, 1989, Asset Pricing in Partially Segmented Markets: Evidence from the Finnish Market, The Journal of Finance, 697-718.

Punom, Chuhan, 1994, Are institutional investors an important source of portfolio investment in emerging market?, The World Bank, International Economics Department, Debt and International Finance Division.

Qian Suna, Wilson H S., Tong, 2000, The effect of market segmentation on stock prices:

The China syndrome, Journal of Banking & Finance, 1875-1902.

René M., Stulz, Walter Wasserfallen, 1995, Foreign equity investment restrictions, capital flight, and shareholder wealth maximization: theory and evidence, The Review of Financial Studies 8(4), 1019–1057.

Stulz R M., Wasserfallen W., 1995, Foreign equity investment restrictions, capital flight, and shareholder wealth maximization: Theory and evidence, The Review of Financial Studies 8(4), 1019-1057.

Sun Q., Tong W H., 1998, The effect of market segmentation on stock prices: The china syndrome, Journal of Banking & Finance 24 (12), 1875-1902.

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

Sun Q., Tong W H., 2000, The effect of market segmentation on stock prices: The China syndrome, Journal of Banking & Finance 24(12), 1875-1902.

Sun Q., Tong W H., 2000, The effect of market segmentation on stock prices: the China syndrome, Journal of Banking & Finance 24(12), 1875-1902.

S Chakravarty, A Sarkar, L Wu, 1998, Information asymmetry, market segmentation and the pricing of cross-listed shares: theory and Evidence from Chinese A and B-shares, Journal of International Financial Markets, Institutions and Money 8, 325 – 355.

Tan L., Chiang T C., Mason J R., Nelling E., 2008, Herding behavior in Chinese stock markets: An examination of A and B-shares, Pacific-Basin Finance Journal 16(1-2), 61-77.

Ting Yang, Sie Ting Lau, 2005, U.S. cross-listing and China's B-share discount, Journal of Multinational Financial Management 15(4-5), 334-353.

Warren Bailey, Julapa Jagtiani, 1994, Foreign ownership restrictions and stock prices in the Thai capital market, Journal of Financial Economics 36, 57-87.

Winnie P H., Poona, Michael Firth, Hung-Gay Fung, 1998, Asset pricing in segmented capital markets: Preliminary evidence from China-domiciled companies, Pacific-Basin Finance Journal, 307-319.

Warren Bailey, Y Peter, Chung, Jun-koo Kang, 2009, Foreign ownership restrictions

and equity price premiums: what drives the demand for cross-border investments, Journal of Financial and Quantitative Analysis.

Wei Sun, 2014, China's A-Share, B-Share, and H-Share Stock Markets and the World Financial Markets: A Cointegration and Causality Analysis, Journal of Applied Business and Economics 16(2), 70 – 80.

Yang J., 2003, Market segmentation and information asymmetry in Chinese stock markets: A VAR analysis, Financial Review 38(4), 591-609.

Yuan D., Zhou X., Li S., 2018, The dynamics of financial market integration between chinese A- and H-Shares, Emerging Markets Finance and Trade 54(13), 2909–2924.

Yin-Hua Yeh, Tsun-siou Lee, Jen-fu Pen, 2002, Stock returns and volatility under market segmentation: the case of Chinese A and B-shares, Review of Quantitative Finance and Accounting 18, 239–257.

Zhou H., 2007, Herding in dual-share stock markets: Evidence from China, Journal of Emerging Markets 12(2), 5.

Zeng W.Y., Luo W.W., 2009, Dynamic conditional correlation among the A-share, B-share and the H-share markets in China, Journal of Business Economics.

Zhao Liuyan, Wang Yiming , 2003, Information flows and volatility spillover across A-share and B-share market in China, Journal of Finance.

Z Jun Lin, Feng Chen, 2005, Value relevance of international accounting standards harmonization: evidence from A- and B-share markets in China, Journal of International Accounting, Auditing and Taxation 14, 79-103

Appendix

a. Data Source Reliability Clarification

In our research for A-H dual-listed companies, we use daily returns, timeframe spanning from 11th October 1990 to 7th February 2020, and unbalanced panel data., but we use only a fraction of it as discussed in previous sections.

Consequently, the data set put into our empirical research consists of 73 companies, thus 73576 observations in total. Our data collection is implemented through two channels: CSMAR and Choice Financial Terminal.

CSMAR is directed by the Wharton Research Data Services, as core Chinese data source. It is widely used as an important reference to top journals in finance academia.

Choice Financial Terminal is widely used terminal for Chinese financial institutions, as Chinese version of Bloomberg/Thomson Reuters.

Official or first-hand data resources provided mainly in Chinese language are referred to for the sake of a cross checking purpose. They are official Chinese financial market information channels, such as Shanghai Stock Exchange, Shenzhen Stock Exchange, National Bureau of Statistics of China, China Securities Regulatory Commission (CSRC); the corresponding media resources designated by CSRC, such as China Securities Journal, Shanghai Securities news, Securities times, Securities Daily and cninfo. World statics are also referred to via World Bank.

b. Elaboration of Model (8.1)

We describe the data we use for the supplementary examination in section seven.

$$Premium_{i,t} = Premium_{i,t-1} + Turnover\ A_{i,t} + Turnover\ H_{i,t} + State\ Ownership_{i,t} + AH\ Outstanding\ Ratio_{i,t} + Asset\ Float_{i,t} + AH\ Volitility\ Ratio_{i,t} + A\ Market\ Index_t + H\ Market\ Index_t + \varepsilon_{i,t}$$

Variable	Description
$Premium_{i,t}$	A-H premium
$Turnover\ A_{i,t}$	Daily turnover of A-share
$Turnover\ H_{i,t}$	Daily turnover of H-share
$State\ Ownership_{i,t}$	Percentage of state ownership on a board basis
$AH\ Outstanding\ Ratio_{i,t}$	Number of outstanding H-shares over A-shares
$Asset\ Float_{i,t}$	Total Value of outstanding A-shares and H-shares
$AH\ Volitility\ Ratio_{i,t}$	Twenty-day volatility of A-share over H-share
$A\ Market\ Index_t$	A-share market index - CSI 300
$H\ Market\ Index_t$	H-share market index - HSI

The turnover variables are retrieved directly from Choice Financial Terminal. State ownership is a very Chinese capital market specific characteristic variable, and is defined as following:

$$State\ Ownership_{i,t} = \frac{Total\ Share_{i,t} - Outstanding\ AShare_{i,t} - Outstanding\ HShare_{i,t}}{Total\ Share_{i,t}}$$

We approximated the state-owned shares via all non-floating percentages, since non-floating shares are either held by SOEs, state institutions, or legal persons subject to the government's guide (Calomiris, Fisman, and Wang 2010). We also neglected the impact from outstanding B-shares since it consists only a tiny portion. On Main Board, only 95 companies, 2.54% of the whole market, issues B-shares and total number of B-shares consists merely 0.00007% of total shares. The number of outstanding H-shares over A-shares is intuitive. Asset float is calculated as the sum of outstanding A-shares and H-shares multiplying their price separately.

$$Asset\ Float_{i,t} = \frac{Asset\ Float\ of\ A_{i,t}\ +\ Asset\ Float\ of\ H_{i,t}}{1000000}$$

wherein:

$$Asset\ Float\ of\ A_{i,t}\ =\ \text{Close price of A}_{i,t}\ \times Outstanding\ AShare_{i,t}$$

$$Asset\ Float\ of\ H_{i,t}\ =\ \text{Close price of H}_{i,t}\ \times Outstanding\ HShare_{i,t}$$

$$1000000 = scaling\ parameter$$

The scaling parameter is only for better data presentation. Twenty-day volatility of A-share over H-share is calculated as below.

$$AH\ Volitility\ Ratio_{i,t} = \frac{\sigma_{20-day\ log\ return\ of\ A_{i,t}}}{\sigma_{20-day\ log\ return\ of\ H_{i,t}}}$$

$$\sigma_{20-day\ log\ return\ of\ A_{i,t}} = \sqrt{\frac{1}{20-1}\sum_{t=0}^{20}(log\ return_{i,t} - \overline{log\ return_{i,t}})^2}$$

$$\sigma_{20-day\ log\ return\ of\ H_{i,t}} = \sqrt{\frac{1}{20-1}\sum_{t=0}^{20}(log\ return_{i,t} - \overline{log\ return_{i,t}})^2}$$

We used twenty as the window length, for there are roughly 20 trading days in a month. We used logarithm return with the Euler's number as the base, for returns are close to log normal distribution.

We describe the data source for this data sample separately. Data related to A-shares and H-shares, including turnover, volatility, share outstanding, and stock price, is from Choice Financial Terminal. Market index data is from Choice Financial Terminal as well. Raw data is processed by us.

c. Original Regression Results

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                Premium   R-squared:                        0.9967
Estimator:                   PanelOLS   R-squared (Between):              0.9987
No. Observations:               73576   R-squared (Within):               0.9935
Date:                Mon, May 04 2020   R-squared (Overall):              0.9967
Time:                        23:03:32   Log-likelihood                 1.276e+05
Cov. Estimator:            Unadjusted
                                        F-statistic:                   7.348e+06
Entities:                        1134   P-value                           0.0000
Avg Obs:                       64.882   Distribution:                 F(3,73573)
Min Obs:                       56.000
Max Obs:                       73.000   F-statistic (robust):          7.348e+06
                                        P-value                           0.0000
Time periods:                      73   Distribution:                 F(3,73573)
Avg Obs:                       1007.9
Min Obs:                       138.00
Max Obs:                       1134.0

                              Parameter Estimates
==============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
Optimistic    3.56e-09   1.29e-09     2.7607     0.0058   1.033e-09   6.088e-09
Pessimistic -1.992e-09  1.568e-09    -1.2707     0.2038  -5.064e-09    1.08e-09
LagPremium      0.9974     0.0002     4021.7     0.0000      0.9969      0.9979
==============================================================================
```

Model (5.1)

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:         NextVolatility   R-squared:                       0.9431
Estimator:                   PanelOLS   R-squared (Between):             0.9439
No. Observations:               75764   R-squared (Within):            2.22e-16
Date:                Mon, May 04 2020   R-squared (Overall):             0.9431
Time:                        23:03:46   Log-likelihood                3.993e+04
Cov. Estimator:             Unadjusted
                                        F-statistic:                  6.276e+05
Entities:                        1107   P-value                          0.0000
Avg Obs:                       68.441   Distribution:                F(2,75762)
Min Obs:                       60.000
Max Obs:                       73.000   F-statistic (robust):         6.276e+05
                                        P-value                          0.0000
Time periods:                      73   Distribution:                F(2,75762)
Avg Obs:                       1037.9
Min Obs:                       124.00
Max Obs:                       1107.0

                              Parameter Estimates
================================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
--------------------------------------------------------------------------------
Optimistic   2.26e-08  8.701e-10     25.971     0.0000   2.089e-08    2.43e-08
Volatility     0.9548     0.0011     870.69     0.0000      0.9526      0.9569
================================================================================
```

Model (5.4)

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:                Premium   R-squared:                       0.9966
Estimator:                   PanelOLS   R-squared (Between):             0.9987
No. Observations:               71731   R-squared (Within):              0.9934
Date:                Mon, May 04 2020   R-squared (Overall):             0.9966
Time:                        23:03:56   Log-likelihood                1.239e+05
Cov. Estimator:             Unadjusted
                                        F-statistic:                  1.059e+07
Entities:                        1107   P-value                          0.0000
Avg Obs:                       64.798   Distribution:                F(2,71729)
Min Obs:                       56.000
Max Obs:                       73.000   F-statistic (robust):         1.059e+07
                                        P-value                          0.0000
Time periods:                      73   Distribution:                F(2,71729)
Avg Obs:                       982.62
Min Obs:                       123.00
Max Obs:                       1107.0

                                 Parameter Estimates
==================================================================================
                     Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
----------------------------------------------------------------------------------
PredictedVolatility     0.0024     0.0004     5.9497     0.0000      0.0016      0.0032
LagPremium              0.9971     0.0003     3182.7     0.0000      0.9964      0.9977
==================================================================================
```

Model (5.6)

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:             AH_Premium   R-squared:                        0.9973
Estimator:                   PanelOLS   R-squared (Between):              0.9995
No. Observations:                 7253   R-squared (Within):               0.9942
Date:                Mon, May 04 2020   R-squared (Overall):              0.9973
Time:                        23:15:09   Log-likelihood                 1.384e+04
Cov. Estimator:             Unadjusted
                                        F-statistic:                    3.01e+05
Entities:                         229   P-value                           0.0000
Avg Obs:                       31.672   Distribution:                  F(9,7244)
Min Obs:                       18.000
Max Obs:                       48.000   F-statistic (robust):           3.01e+05
                                        P-value                           0.0000
Time periods:                      73   Distribution:                  F(9,7244)
Avg Obs:                       99.356
Min Obs:                       3.0000
Max Obs:                       182.00

                              Parameter Estimates
================================================================================
                Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
--------------------------------------------------------------------------------
Lag_Premium        0.9990     0.0009     1082.6     0.0000      0.9972      1.0009
Turnover_A         0.1708     0.0283     6.0294     0.0000      0.1152      0.2263
Turnover_H         0.1876     0.0559     3.3561     0.0008      0.0780      0.2972
State_Own         -0.0129     0.0041    -3.1266     0.0018     -0.0210     -0.0048
HOut_Over_AOut    -0.0049     0.0027    -1.7854     0.0742     -0.0102      0.0005
AH_Folat        8.068e-10  1.281e-09     0.6299     0.5288  -1.704e-09   3.318e-09
Sigma_AoverH       0.0013     0.0010     1.3417     0.1798     -0.0006      0.0033
CSI                0.0081     0.0004     20.917     0.0000      0.0074      0.0089
HSI               -0.0092     0.0005    -18.747     0.0000     -0.0102     -0.0083
================================================================================
```

Model (8.1)

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:             AH_Premium   R-squared:                       0.5658
Estimator:                   PanelOLS   R-squared (Between):             0.9105
No. Observations:                 7253   R-squared (Within):              0.0622
Date:                Wed, May 06 2020   R-squared (Overall):             0.5658
Time:                        21:31:26   Log-likelihood                  -4629.9
Cov. Estimator:            Unadjusted
                                        F-statistic:                     1180.2
Entities:                         229   P-value                          0.0000
Avg Obs:                       31.672   Distribution:                 F(8,7245)
Min Obs:                       18.000
Max Obs:                       48.000   F-statistic (robust):            1180.2
                                        P-value                          0.0000
Time periods:                      73   Distribution:                 F(8,7245)
Avg Obs:                       99.356
Min Obs:                       3.0000
Max Obs:                       182.00

                              Parameter Estimates
================================================================================
                 Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
--------------------------------------------------------------------------------
Turnover_A        -2.5386     0.3599    -7.0537     0.0000     -3.2441     -1.8331
Turnover_H         6.8818     0.7088     9.7090     0.0000      5.4924      8.2713
State_Own          0.4212     0.0524     8.0404     0.0000      0.3185      0.5239
HOut_Over_AOut     0.8475     0.0333     25.452     0.0000      0.7823      0.9128
AH_Folat       -1.958e-07  1.618e-08    -12.104     0.0000  -2.275e-07  -1.641e-07
Sigma_AoverH       0.2921     0.0122     23.996     0.0000      0.2683      0.3160
CSI                0.0011     0.0050     0.2198     0.8260     -0.0086      0.0108
HSI               -0.0075     0.0063    -1.1889     0.2345     -0.0198      0.0048
================================================================================
```

Only 57% of variation can be explained

d. Further Issues of Share Classification in Chinese Stock Market

Chinese companies registered in mainland China and being traded in either mainland or Hong Kong, or even both, follow its specific equity ownership binary plan – equity structure embracing both tradable and non-tradable parts. Non-tradable parts are state-owned shares and State-owned legal person shares, which can only be traded by qualified identities via auction or transfer agreement under the direct supervision from CSRC. Before the split structure reform, due to aforementioned specific ownership status, the non-tradable portion consisting more than one-third of the domestically listed companies remained as one of main reasons for pricing unfairness. It was because remaining illiquid portion until the split-share structure reform started in 2005 and later eased the percentage gradually. The process is nearing its end.

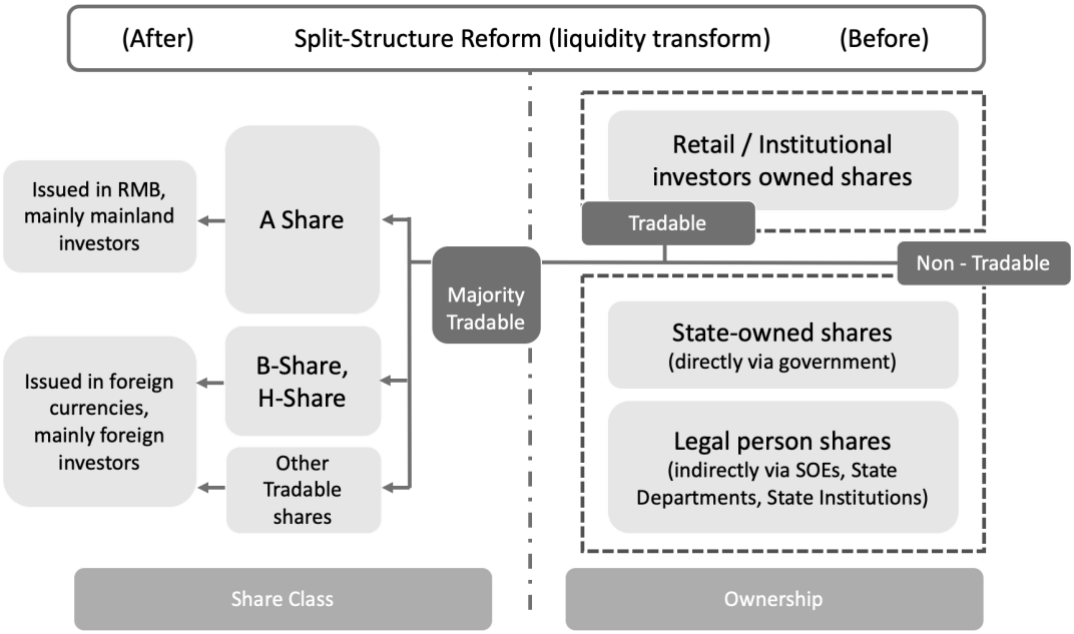

Table. Appendix.4

Tradable portion are allowed for investors in secondary markets. They are A-share, H-share, B-shares who are differentiated via trading currencies, listed markets and ownership qualification restriction, if applicable, such as nationality, identity or ownership upper limits. At times, tradable parts face restriction period and are not fully

transferable for a certain period of time. However, the measure reflects, most of the times, corporate strategies such as managerial or employer incentive schemes or sub-period effect such as newly closed M&A deals, rather than the ownership issues. As of April 2020, tradable and non-restricted shares listed in Shanghai and Shenzhen Stock Exchanges consists 88% and 87% correspondingly.

Pricing mechanism for non-tradable parts are entirely deviated from market pricing mechanism. Instead, it is subject to national regulations or Company Laws describing OTC trading pricing rules that are applicable to specific legal entities. Thus, in Chinese security markets, the characteristics of the tradable part stocks, consisting rather high portion of total, are good representatives for Chinese company investors' specific trading behaviors and thus phenomena. In contrast, the non-tradable portion can be of a symbolic feature for well-known Chinese state-ownership issue.