Investor Sentiment, Idiosyncratic Risk, and Stock Price Premium: Evidence From Chinese Cross-Listed Companies

Yuan Li¹,² and Yu Zhang³

Abstract
We investigate whether idiosyncratic risk and investor sentiment play important roles in the price disparity between China A-shares and H-shares. To build the investor sentiment indices and decompose them into different fragments for both markets, we use both principal component analysis (PCA) and partial least squares (PLS) approaches. We further look at how idiosyncratic risk affects stock mispricing and how it deals with investor sentiment. We find that the price premium of A-shares over H-shares is strongly linked to the sentiment differential. We also discover that idiosyncratic risk has a major effect on the price premium of cross-listed companies. Moreover, a larger sentiment differential reinforces the impact of idiosyncratic risk on the price disparity. The above results remain robust after controlling for other economic factors.

Keywords
price premium, investor sentiment, idiosyncratic risk, cross-listed

Introduction
International diversity attracts free capital moving across borders, and numerous companies have gained capitals outside of their home countries over the last few decades. Cross-listing refers to a company listing its shares in the overseas market as well as its home market. The emphasis of this article is on Chinese cross-listed companies that issue A-shares and H-shares. There are some distinctions between the two groups of stocks. A-shares are limited to domestic investors and are listed and exchanged on the stock exchanges in Mainland China. H-shares, on the other hand, are listed on the stock exchange in Hong Kong and have their headquarters in Mainland China. It is worth noting that H-shares are available for purchase by Hong Kong and international buyers.

If stock markets are effective and arbitrage is easily implemented, foreign and domestic securities should have the same price. However, according to Fernald and Rogers (2002), A-shares in Mainland China are traded at a premium to H-shares. Although the price disparity of cross-listed companies has been discussed for years, there has been comparatively limited literature on both investor sentiment and idiosyncratic risk. Previous research has verified the individual impact of market sentiment or idiosyncratic risk on the mispricing, such as Baker and Stein (2004), Baker and Wurgler (2006), Pontiff (2006), Schmeling (2009), and Beckmann et al. (2015). The combined effect of investor sentiment and idiosyncratic risk on stock mispricing is the subject of this article. We look into how idiosyncratic risk influences price disparities as investor sentiment shifts. Another distinction from previous research is that this article uses two methods, principal component analysis (PCA) and partial least squares (PLS) approaches, to create investor sentiment measures. The following questions are addressed in this article: How much does the difference in investor sentiment between the two markets, whether total or local, impact the price gap between A-shares and H-shares? What role does idiosyncratic risk play in the stock price disparity between cross-listed companies? What is the relationship between idiosyncratic risk and investor sentiment?

This article makes several contributes to the current literature. First, voluminous research on the price disparity of cross-listed firms has been conducted, with an emphasis on developed markets. Markets we investigate in this study concentrate on emerging markets with a limited degree of openness. We apply a panel data methodology to assess the degree...
to which idiosyncratic risk and investor sentiment explain the price disparity.

Second, it is an innovative exercise to use both PCA and PLS methods to build investor sentiment to investigate price deviations in cross-listed securities. Our article is unique in that it decomposes investor sentiment into various fragments and tests the impact of total and local sentiment differentials, as well as idiosyncratic risk, on mispricing. Previous research has always used a single index to assess investor sentiment. For example, to measure investor sentiment, Arquette et al. (2008) use the price–earnings ratio as a metric to examine the price disparity of Chinese cross-listed companies. We have some reservations about this metric since a single index does not necessarily provide comprehensive economic viewpoints. Other studies using the composite sentiment index often apply one method, either PCA method (Baker et al., 2012; Fu et al., 2020) or PLS method (Wu et al., 2017). This article, as stated above, employs both PCA and PLS methods to construct the composite sentiment indices and test their roles in the price disparity between China A-shares and H-shares.

Finally, the interaction effect between investor sentiment and idiosyncratic risk is investigated in this article. Gagnon and Karolyi (2010) examine arbitrage risk using idiosyncratic volatility and Stambaugh et al. (2015) follow suits. In this article, we test the marginal effect of idiosyncratic risk on the price deviation of A-shares and H-shares while sentiment is changing. Prior studies have confirmed that differences in demand, liquidity, risk, and information asymmetry are associated with the underlying stock price premium (Bailey & Jagtiani, 1994; Chakravarty et al., 1998; Chan et al., 2008; Grossmann et al., 2007; Karolyi et al., 2009; Mei et al., 2005; Rui et al., 2007; Stulz, 1981; Stulz & Wasserfallen, 1995; Sun & Tong, 2000; Wang & Jiang, 2004). After controlling for these traditional factors, according to our findings, a high sentiment differential reflects the changes of market sentiment to examine price variations. The price–earnings ratio is used to gauge investor sentiment. Other studies, such as Guo (2013), also investigate this mispricing and adopt the same single index to reflect the changes of market sentiment to examine price disparities in Chinese cross-listed firms. Compared with the above studies, we divide the sentiment indices into total, local, and global fragments using both PCA and PLS approaches. Unlike Baker et al. (2012), who examine market deviations for just three dual-listed firms, we look at a much larger number of Chinese firms. When market investor sentiment is relatively high, according to Wu et al. (2017), investors tend to be more positive about the future and buy more stocks. As a result, this market’s stock prices rise. We conjecture that the price disparity of cross-listed companies is positively linked to the sentiment differential. We thereby hypothesize the following:

**Hypothesis 1:** The price disparity between A-shares in Mainland China and H-shares in Hong Kong is positively related to the total (local) sentiment differential.
Idiosyncratic Risk and Mispricing of Chinese Stocks

In the following sections, we also test how idiosyncratic risk impacts cross-listed company mispricing. Idiosyncratic risk is the element of the fundamental value and is uncorrelated to the returns of other assets that are available to the arbitrageur (Pontiff, 2006). Merton (1987) observes that when stock markets are segmented, the idiosyncratic risk would increase potential expected returns. He also claims that investors would pay a fixed fee to become informed. Shleifer and Vishny (1997), on the other hand, contend that stocks are not priced rationally, and stocks with high idiosyncratic variance are likely to be overpriced. Idiosyncratic risk deters arbitrage. According to Pontiff (2006), the single greatest cost faced by arbitrageurs is the idiosyncratic risk. The above studies investigate the relationship between idiosyncratic risk and stock price. Afterward, using 12 cross-listed firms from 1980 to 2002, De Jong et al. (2009) leverage the price deviation. They suggest that the arbitrage of cross-listed companies is affected by idiosyncratic return volatility and negative returns are likely to impede arbitrage. According to Gagnon and Karolyi (2010), increases in price disparities are positively correlated with proxies reflecting holding costs that would hinder arbitrage. They argue that idiosyncratic risk is the most significant statistical and economic proxy for holding costs. Similar work are Beckmann et al. (2015), Stambaugh et al. (2015), Wu et al. (2017), and Li and Ran (2020). All the above studies do not emphasize the mispricing of stocks from the perspective of idiosyncratic risk as well as investor sentiment for cross-listed companies. In the parts that follow, we examine how idiosyncratic risk impacts the price variation of stocks of cross-listed companies with varying investor sentiment. When investor sentiment becomes high, investors appear to become hesitant to arbitrage since holding costs would rise (Pontiff, 2006).

**Hypothesis 2:** The price deviations between A-shares and H-shares are positively linked to idiosyncratic risk.

**Hypothesis 3:** Investor sentiment reinforces the connection between idiosyncratic risk and price deviations between A-shares and H-shares.

Data Description and Empirical Design

Basic Approach

In building the stock market sentiment index, we use both PCA and PLS approaches. To derive the first principal component, Baker and Wurgler (2006) use six proxies to represent market sentiment. PCA approach is used to figure out data’s underlying structure. Baker and Wurgler (2007), Stambaugh et al. (2015), and Wu et al. (2017) employ the same method to prove its validation and test their hypotheses. PCA procedure is as follows. The eigenvalue and eigenvector are calculated after standardizing the sentiment proxies. The composite index is then generated by combining the proxies in a linear fashion using the eigenvector that best represents the total variance.

To predict the aggregate stock market, Huang et al. (2015) use PLS approach instead of PCA, to propose a new aligned investor sentiment index. They argue that, according to Baker and Wurgler (2006, 2007), investor sentiment cannot be distinguished from the common approximation error component of the proxies that are unrelated to returns. To resolve this challenge, they filter out the irrelevant components in sentiment proxies, as Wold (1966, 1975) did, to build a new, better predictor. To construct the PLS index, first, we run N time-series regressions as follows:

\[ x_{i,t=1} = \alpha_{i,0} + \alpha_i R_t + u_{i,t-1}, \]

where \( x_i \) is the investor sentiment proxy, \( R_t \) is the realized stock returns. The loading \( \alpha_i \) captures the sentiment proxy’s sensitivity to investor sentiment as measured by future stock return \( R_t \).

Second, run T cross-sectional regressions as follows:

\[ x_{i,t} = \beta_i + S_{t,PLS} \tilde{\alpha}_i + v_{i,t-1}, \]

where \( \tilde{\alpha}_i \) is the loading estimated in the first regression and \( S_{t,PLS} \) is the estimated investor sentiment. PLS index will reduce the noise asymptotically and have a better forecasting efficiency, according to theoretical and empirical findings in Kelly and Pruitt (2013, 2015). Therefore, findings in Huang et al. (2015) confirm that PLS index outperforms PCA index with a common noise component for the US stock market. Wu et al. (2017) also test the validity of the method by examining the mispricing of American Depository Receipt (ADR). Against this backdrop, to properly shape the valid market-based sentiment indices, our investigation employs both PCA and PLS approaches. We compare these two methods in the investigation of the price disparity of cross-listed companies in China.

Data and Sample Periods

This article relies on data from the CEIC and Wind databases. From January 2003 to December 2019, our sample includes A-shares and H-shares that are listed on the Shanghai and Hong Kong stock exchanges, respectively. There are 95 companies issuing both A-shares and H-shares by the end of the year 2019. To ensure that we have adequate data for all the variables, 75 companies are excluded because A-shares or H-shares of the companies were listed after 2003. Finally, we get 20 companies whose listing date is before 2003 on both markets. To calculate market returns, the Shanghai
Composite Indices and the Hang Seng Index are used. The risk-free rate is subtracted from the market return to measure the excess market return. The risk-free rate is described as the monthly rate of a 1-year time deposit in Mainland China, according to the convention. The risk-free rate in Hong Kong is the Hong Kong Interbank Offered Rates (3-month HIBOR). We also calculate market returns using value-weighted market indices and obtain similar results.

Investor Sentiment
Following Verma and Soydemir (2009), Baker et al. (2012), Han and Li (2017), Wu et al. (2017), and Li and Ran (2020), the market turnover ratio (aggregate market value-weighted), the price–earnings ratio, and the total number of initial public offerings (IPOs) are used as substitutes to construct the total sentiment index for the stock market in Hong Kong. For stock market in Mainland China, however, we combine the market turnover ratio (aggregate market value-weighted), the price–earnings ratio, and the newly opened accounts number of individual investors (in Shanghai Stock Exchange) to calculate the total sentiment index. It is worth noting that in the absence of IPO data, we use the number of new accounts instead of IPO data for the Shanghai Stock Exchange. The use of the above sentiment proxies in our article is motivated primarily by data availability and economic considerations, and the details of the proxies can be found in Li and Ran (2020).

To create a “total” investor sentiment for the Mainland China and Hong Kong stock markets, we use both PCA and PLS approaches described above. Global and local investor indices are then established. The methods for calculating global and local sentiment indices can be found in Baker et al. (2012).

Idiosyncratic Risk
According to Pontiff (2006), the single greatest cost faced by arbitrageurs is the idiosyncratic risk. As is pointed out, idiosyncratic risk is a component of intrinsic valuation that is irrelevant to market returns or the returns of some other hedge portfolios. Gagnon and Karolyi (2010), as well as Wu et al. (2017), investigate how idiosyncratic risk affects cross-listed markets. Following Gagnon and Karolyi (2010), we estimate a regression of the daily return differential using contemporaneous leading, lagged daily stock index returns from Mainland China and Hong Kong, and the currency shifts, within each company quarter:

\[
R_{A-H,t} = \alpha + \sum_{i=1}^{3} \beta_{1}^{A} R_{M,t-1}^{A_{i}} + \sum_{i=1}^{3} \beta_{1}^{H} R_{M,t-1}^{H_{i}} + \sum_{i=1}^{3} \beta_{1}^{FX} R_{F_{X,t-1}} + e_{A-H,t},
\]

where \(R_{A-H,t}\) is the return difference between A-shares and H-shares. \(R_{M,t-1}^{A_{i}}\) and \(R_{M,t-1}^{H_{i}}\) are A-shares and H-shares stock market index returns. \(R_{F_{X,t-1}}\) is the currency return of Hong Kong dollars relative to the Chinese Yuan. The idiosyncratic risk is calculated as the standard deviation of the residuals of Equation 3. As is pointed out in Gagnon and Karolyi (2010), due to the sluggish spread of information, the stock price show a delayed reaction to market returns. In our model, we regress the return differential to the contemporaneous market and currency returns, as well as 1-day leading and lagging market and currency returns.

Table 1 provides the summary statistics of this article.

Control Variables
Our research links investor sentiment and idiosyncratic risk to the price disparity found in Chinese cross-listed companies. Instead of using a single sentiment index, we explore the relationship among investor sentiment, idiosyncratic risk, and stock price differentials using both PCA and PLS sentiment indices.

The stock price premium of A-shares over H-shares is defined as the dependent variable in the study. Previous research has looked at the price disparity between cross-listed stock markets and offered several reasons for the price deviations. We add four commonly used variables in the literature, as well as other control variables for reference.

Demand differential. According to studies such as Sun and Tong (2000) and Darrat et al. (2006), \(\text{DEM}_{u}\) stands for the demand differential, which is calculated by dividing the number of outstanding A-shares by the number of outstanding H-shares. We use this proxy as the variable representing the demand differential of the two markets. When the number of tradable A-shares outstanding increases, A-shares price falls and the price difference between A- and H-shares narrows. We conjecture that \(\text{DEM}_{u}\) is inversely linked to the price gap: The larger the number of tradable A-shares, the bigger the \(\text{DEM}_{u}\), and hence the lower the price premium.

Asymmetric information. Previous literature has used market capitalization to quantify asymmetric information. According to Chan and Kwok (2005) and Karolyi et al. (2009), larger corporations with higher market capitalizations often release more information. The total amount of market capitalization of the A-shares and H-shares is used to measure \(\text{INF}_{u}\). We employ \(\text{INF}_{u}\) to quantify the current asymmetric information effects for cross-listed companies. Larger companies with greater information transparency have a smaller price disparity between their cross-listed shares.

Liquidity differential. \(\text{LIQ}_{u}\) represents the relative liquidity of A-shares and H-shares. Since the Amihud ratio (Amihud, 2002) is always used as the proxy for liquidity, we calculate

\[
\text{LIQ}_{u} = \frac{1}{\text{Holo}_{u}} \sum_{t=1}^{T} \frac{|A_{t} - H_{t}|}{A_{t} + H_{t}},
\]
this liquidity ratio in the same way Chan et al. (2008) and Wu et al. (2017) did. A stock’s illiquidity is defined as the average ratio of its daily absolute return to its trading volume in RMB, denoted as $LIQ_{A,t}$ and $LIQ_{H,t}$. We conjecture that if A-shares are comparatively less liquid, it would result in a lower price premium.

**Risk differential.** $RISK_{i,t}$ represents the risk differential as the ratio of A-shares returns standard deviations to H-shares return standard deviations. We compute the rolling standard deviation of the previous 10-day returns, as seen in Schwert (2002). By raising A-shares price, a comparatively high risk of A-shares would result in a higher price premium.

**Others.** We introduce a $DUMMY$, variable to represent the new cross-border investment program in China (Shanghai-Hong Kong Stock Connect). When $DUMMY$ equals to 1, it indicates the period after the launch of the program. In the final regression, we substitute the US investor sentiment index (BW index), for the global sentiment. We also consider the shifts in the exchange rate (RMB/HKD) in our empirical analysis.

**Model Specification**

First, we plan to look at how investor sentiment differential impacts the price disparity between the two types of shares. We analyze how fluctuations in stock price premiums are correlated to market investor sentiment across companies using a panel data approach. The full models are as follows:

$$Prem_{i,t} = \alpha + \beta_1 Prem_{i,t-1} + \mu_1 Controls_{i,t} + u_{i,t}$$

(4)

$$Prem_{i,t} = \alpha + \beta_1 Diff_{-S_{Total,t-1}} + \delta_1 Prem_{i,t-1} + \mu_1 Controls_{i,t} + u_{i,t}$$

(5)

where $Prem_{i,t}$ is A-shares price premium in RMB. $Diff_{-S_{Total,t-1}}$ is the difference of total (local) sentiment between A-shares and H-shares markets ($Diff_{-S_{Total,t-1}} = S_{Total,t-1} - S_{Total_{HR,t-1}}$), and $S_{Global,t}$ is global investor sentiment. $Controls_{i,t}$ include $DEM_{i,t}$, $INF_{i,t}$, $LIQA_{i,t}$, $LIQH_{i,t}$, $RISK_{i,t}$, $DUMMY$, and $EX_{i,t}$ equals to $PCA$ or $PLS$.

As mentioned in prior studies, investor sentiment would influence the mispricing indirectly by idiosyncratic risk (Wu et al., 2017). In the second step, we intend to investigate whether the magnitude of the price premium varies over time in a way that is related to the idiosyncratic risk. We conduct the following regressions to examine the magnitude of the price premium that arises from idiosyncratic risk, with and without sentiment controls. We add the interaction term of idiosyncratic risk and investor sentiment to explore the indirect impact of investor sentiment. The models are as follows:
premium. The sign of coefficient of that comparatively illiquid A-shares will have a relatively consistent with the differential liquidity theorem, which states Illiquid shares will sell at a lower price compared with liquid down due to the relative illiquidity of the A-shares. cant at 1% level. It suggested that the price gap would narrow across-listed companies since their coefficients are signif-

\[
\text{Prem}_{it} = \alpha + \beta_1 \text{IVOL}_{it} + \beta_2 \text{Diff}_{St-1} + \beta_3 \text{IVOL}_{it},
\]

where \( \text{Prem}_{it} \) is A-shares price premium expressed in RMB. \( \text{Diff}_{St-1} \), \( \text{IVOL}_{it} \), and \( \text{St}_{Global} \) are defined as above. \( \text{IVOL}_{it} \) is the idiosyncratic risk for each company.

Empirical Results

Sentiment Differential and Price Disparity Between A-Shares and H-Shares

Table 2 provides empirical results for the impact of sentiment differential on price disparity between A-shares and H-shares. Positive coefficients in the table indicate that the increase of the variables would lead to a larger price premium. The findings in Column (1) of Panel A reveal that by using PCA approach, the difference in investor sentiment between the two markets plays a critical role in deciding the price disparity of cross-listed stocks. The coefficient of \( \text{Diff}_{St-1} \) is 0.0044 and significantly positive at 1% level. The result indicates that the sentiment differential is a substantial factor explaining the stock price premium. Columns (2) allows us to examine the effect of \( \text{Diff}_{St-1} \) with control variables included. After controlling for a variety of variables, \( \text{DEMA}, \text{INFR}, \text{LIQA}_{it}, \text{LIQH}_{it}, \text{RISK}_{it}, \) and \( \text{EX} \), the results are still robust and sentiment differential is favorably associated with a stock price premium. As to control variables, the sign of \( \text{DEMA} \) is significantly negative at 1% level. The A-shares price would fall as the number of tradable A-shares increases. Under this specification, the coefficient of \( \text{INFR} \) is insignificantly negative, which is in line with the previous theorem on information asymmetry. The result clearly states that the larger the overall market capitalization, the closer A- and H-shares prices are. \( \text{LIQA}_{it} \) and \( \text{LIQH}_{it} \) are both closely involved with the price gap across cross-listed companies since their coefficients are significant at 1% level. It suggested that the price gap would narrow down due to the relative illiquidity of the A-shares. Illiquid shares will sell at a lower price compared with liquid shares, resulting in a smaller price gap. The result is consistent with the differential liquidity theorem, which states that comparatively illiquid A-shares will have a relatively lower price than H-shares, leading to a lower A-shares price premium. The sign of the coefficient of \( \text{RISK}_{it} \) is also significantly positive at 1% level. According to the risk hypothesis mentioned above, A-shares with relatively high risk would increase in price, resulting in a larger A-shares price premium. What’s more, the coefficient of \( \text{EX} \) is negative, which indicates the price deviation between the two types of shares would be mitigated by the depreciation of RMB. To summarize, the benchmark model findings validate our crucial findings. In Columns (3) and (4), we replace \( \text{Diff}_{St-1} \) with \( \text{Diff}_{St-1} \) and the results suggest that after taking into account the control variables, \( \text{Diff}_{St-1} \) is still favorably correlated with the price disparity between cross-listed companies. When control variables are included, \( \text{Diff}_{St-1} \) is still significantly positive at 1% level. Results of the control variables are robust and generate quite similar results as above. The introduction of \( \text{St}_{Global} \) has little impact on our results either. It is quite interesting that the coefficients of \( \text{DEMY}_{it} \) are positive in Columns (7) and (8). It seems that the launch of Shanghai-Hong Kong Stock Connect does not narrow the price disparity as expected. Part of the reason, in our view, is that A-shares market is more volatile and easier to speculate. After the launch of the program, investors could invest in the Mainland China market. The situation of overvaluing in A-shares price deteriorates, and thereby increases the price premium. Results from Panel A indicate clearly that the difference in the sentiment is a major factor in deciding the price disparity between the two types of shares. Results presented above are consistent with those of previous research such as Beckmann et al. (2015) and Wu et. al (2017).

Findings in Panel B are quite close to the results in Panel A. We can conclude from Column (1) that sentiment differential is strongly linked to the A-H stock price gap. The coefficient of \( \text{Diff}_{St-1} \) is 0.0078 and significantly positive at 1% level. We can conclude that the greater the sentiment differential, the larger the stock price disparity. The introduction of control variables does not change the results, and \( \text{Diff}_{St-1} \) is still the influential factor in determining the price premium. The signs of control variables are identical to those results in Panel A. The coefficient of \( \text{Diff}_{St-1} \) is significantly positive at 1% level. The above results remain robust with the introduction of \( \text{St}_{Global} \). We may assume from the preceding debate that, by using PLS approach, there is a favorable association between stock price differentials and investor sentiment in cross-listed firms. This is consistent with the results when PCA method is adopted. To sum it up, for Chinese cross-listed A-shares and H-shares, investor sentiment differential plays a major role in the price deviation between the two types of shares. Different from Huang et al. (2015), there is no big difference between the two methods when examining the Chinese market.

Idiosyncratic Risk and Price Disparity Between A-Shares and H-Shares

Idiosyncratic risk, as stated in previous studies, will increase the price disparity between cross-listed shares. We plan to investigate the primary factors that influence the price gap. We would test how idiosyncratic risk behaves when investor
Table 2. Time Series Regressions for Cross-Listed Companies.

| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|
| Panel A: PCA method | | | | | | | | |
| \( \text{Diff} \_S_{\text{Total}}^{\text{St}} \) | 0.0044*** | 0.0037*** | 0.0043*** | 0.0043*** | 0.0043*** | 0.0043*** | 0.0043*** | 0.0043*** |
| \( \text{Diff} \_\_S_{\text{Local}}^{\text{St}} \) | 0.0041*** | 0.0034*** | 0.0041*** | 0.0041*** | 0.0041*** | 0.0041*** | 0.0041*** | 0.0041*** |
| \( \_S_{\text{Global}}^{\text{St}} \) | -0.0044 | -0.0047* | -0.0047* | -0.0047* | -0.0047* | -0.0047* | -0.0047* | -0.0047* |
| \( \text{DEMM}_{it} \), \( \text{INF}_{it} \), \( \text{LIQA}_{it} \), \( \text{LIQH}_{it} \), \( \text{RISK}_{it} \), \( \text{EX}_{it} \), and \( \text{DUMMY}_{it} \) | 0.0092** | 0.0093** | 0.0092** | 0.0093** | 0.0092** | 0.0093** | 0.0092** | 0.0093** |
| Panel B: PLS method | | | | | | | | |
| \( \text{Diff} \_S_{\text{Total}}^{\text{St}} \) | 0.0078*** | 0.0066*** | 0.0078*** | 0.0078*** | 0.0078*** | 0.0078*** | 0.0078*** | 0.0078*** |
| \( \text{Diff} \_\_S_{\text{Local}}^{\text{St}} \) | 0.0077*** | 0.0063*** | 0.0077*** | 0.0077*** | 0.0077*** | 0.0077*** | 0.0077*** | 0.0077*** |
| \( \_S_{\text{Global}}^{\text{St}} \) | -0.0046* | -0.0049* | -0.0049* | -0.0049* | -0.0049* | -0.0049* | -0.0049* | -0.0049* |
| \( \text{DEMM}_{it} \), \( \text{INF}_{it} \), \( \text{LIQA}_{it} \), \( \text{LIQH}_{it} \), \( \text{RISK}_{it} \), \( \text{EX}_{it} \), and \( \text{DUMMY}_{it} \) | 0.0092** | 0.0093** | 0.0092** | 0.0093** | 0.0092** | 0.0093** | 0.0092** | 0.0093** |

Note. The stock price of A-shares over H-shares, expressed in RMB, is the dependent variable in Panels A and B. Global sentiment (\( S_{\text{Global}}^{\text{St}} \)) is identified as the component of the total sentiment (\( S_{\text{Total}}^{\text{St}} \)). Local sentiment (\( S_{\text{Local}}^{\text{St}} \)) is the residual of the regression: \( S_{\text{Total}}^{\text{St}} = \alpha + \beta S_{\text{Global}}^{\text{St}} + S_{\text{Local}}^{\text{St}} \), for each market. \( \text{Diff} \_S_{\text{Total}}^{\text{St}} = S_{\text{Total}}^{\text{St}} - S_{\text{Total}}^{\text{St}} \_t-1 \), \( \text{Diff} \_S_{\text{Local}}^{\text{St}} = S_{\text{Local}}^{\text{St}} - S_{\text{Local}}^{\text{St}} \_t-1 \), \( \text{DEMM}_{it} \), \( \text{INF}_{it} \), \( \text{LIQA}_{it} \), \( \text{LIQH}_{it} \), \( \text{RISK}_{it} \), \( \text{EX}_{it} \), and \( \text{DUMMY}_{it} \) are the above-mentioned control variables. PCA and PLS approaches are used to construct the sentiment indices in Panels A and B, respectively. PCA = principal component analysis; PLS = partial least squares. Samples include observations from January 2003 to December 2019. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

sentiment changes. To begin, we sort the stock price premium between the two types of shares based on investor sentiment indices and company-level idiosyncratic risk. Table 3 shows the simple two-way sorting. On one hand, as idiosyncratic risk grows, the price gap of A-shares compared with H-shares rises in all four of our categories.
Regardless of the size of the sentiment differential, the price premium is higher in high-idiosyncratic-risk categories. On the other hand, in most situations, the price premium in small-sentiment-differential categories is relatively small on average. We may infer that the price premium in large-sentiment-differential groups is higher on average than the price premium in small-sentiment-differential groups. Based on this, we will explore whether the difference in investor sentiment has an indirect impact on stock price premium through idiosyncratic risk.

We run several regressions according to Equations 6 and 7. As seen in Table 4, idiosyncratic risk is a statistically important element in the stock price premium of A-shares over H-shares. What should be pointed out is that data for the BW index are only updated to the end of 2018 (http://people.stern.nyu.edu/jwurgler/), so the sample period in Table 4 is from January 2003 to December 2018. The coefficient of $IVOL_{it}$ in Column (1) of Panel A indicates a significantly positive relationship between price gap and idiosyncratic risk. The coefficient of $IVOL_{it}$ is 0.5283 and significantly positive at 1% level. The stock price premium of A-shares over H-shares rises as idiosyncratic risk increases and is in line with the findings in Table 3. In Column (2), investor sentiment is favorably linked to the price premium, but the coefficient is much smaller than the idiosyncratic risk in absolute value. The regression slope of idiosyncratic risk in absolute value. The regression slope of idiosyncratic risk is positive in the next column, implying that a higher sentiment differential would have a larger marginal effect on the price premium through idiosyncratic risk. $Diff_{S_{Total},-1}$ becomes insignificantly negative when the interaction term ($IVOL_{it} \times Diff_{S_{Total},-1}$) is included. It indicates that the effect of $Diff_{S_{Total},-1}$ is overshadowed by $IVOL_{it} \times Diff_{S_{Total},-1}$. When local sentiment is used instead of total sentiment in Columns (4) to (7), the findings are quite close to those obtained from the total sentiment. On one hand, $Diff_{Local,1}$ by itself is linked to the price disparity positively. When the interaction term is included, however, the coefficient of $Diff_{Local,1}$ becomes negative. Since the interaction term is still significantly positive, it plays a bigger role than $Diff_{Local,1}$. When control variables are included, the sign of the coefficient of idiosyncratic risk remains the same. Results are robust when we replace $S_{Total}^{Global}$ with $S_{Total}^{Local}$. We may infer from Panel A that idiosyncratic risk is favorably linked to the price premium between the two types of shares. A higher total (local) sentiment differential would reinforce the sensitivity of idiosyncratic risk to the stock price premium.

Panel B presents the findings while sentiment indices are calculated using PLS method. In all regressions, the idiosyncratic risk in Panel B is positively linked to the stock price premium. When control variables are allowed, the results are robust. The interaction term ($IVOL_{it} \times Diff_{S_{Total},-1}$) remains significantly positive, as seen in Panel A, indicating that a higher total sentiment differential reinforces the vulnerability of idiosyncratic risk to the stock price premium. As we replace the total sentiment differential with the local sentiment differential, the coefficient of the interaction term ($IVOL_{it} \times Diff_{Local,1}$) is still positive and significant. What should be noted is that when the interaction term is included in Panel B, the sentiment differential term, $Diff_{Total,1}$ (or $Diff_{Local,1}$), does not play many roles. The coefficient of $Diff_{Total,1}$ (or $Diff_{Local,1}$) becomes negative and insignificant, suggesting the dominating role of the interaction term. Control variables show similar characteristics as in Table 2. In all of our regressions, the dummy variable representing Shanghai-Hong Kong Stock Connect is strongly positive. Obviously, the launch of the program does not reduce the price disparity as expected. To
| Variable                        | (1)      | (2)      | (3)      | (4)      | (5)      | (6)      | (7)      | (8)      | (9)      | (10)     | (11)     | (12)     | (13)     | (14)     | (15)     |
|--------------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| **Panel A: PCA method**        |          |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| $IVOL_{it}$                    | 0.5283***| 0.4672***| 0.4207***| 0.4767***| 0.4296** | 0.4616***| 0.4146** | 0.5827***| 0.5763***| 0.4115** | 0.3546** | 0.3639** | 0.4361***| 0.3898** | 0.3026*  |
|                                | (3.179)  | (2.795)  | (2.508)  | (2.855)  | (2.566)  | (2.760)  | (2.472)  | (3.350)  | (3.110)  | (2.295)  | (2.019)  | (2.578)  | (2.298)  | (1.670)  |          |
| **Diff** $- \sigma_{it}$       | 0.0041***| -0.0055  |          |          |          |          |          |          |          |          |          |          |          |          |          |
|                                | (3.127)  | (-1.538) |          |          |          |          |          |          |          |          |          |          |          |          |          |
| $IVOL_{it} \times \text{Diff}_{it}$ & 0.2966***|          |          |          |          |          |          |          |          |          |          |          |          |          |          |
|                                | (2.891)  |          |          |          |          |          |          |          |          |          |          |          |          |          |          |
| $\Delta_{it}$                  | 0.0039***| -0.0063* | 0.0039***| -0.0063* | 0.0038***|          |          | 0.0033***| 0.0037***| -0.0065* | -0.0070**|          |          |          |          |
|                                | (2.892)  | (-1.789) | (2.917)  | (-1.777) | (2.803)  |          |          | (2.464)  | (2.730)  | (-1.834) | (-2.019) |          |          |          |          |
| $\sigma_{it}$                  | 0.3144***|          | 0.3141***|          |          |          |          | 0.3133***| 0.3147***|          |          |          |          |          |          |
|                                | (3.123)  | (3.120)  |          |          |          |          |          | (3.113)  | (3.161)  |          |          |          |          |          |          |
| $\Delta_{it}$                  | -0.0040  | -0.0040  | -0.0037  |          |          |          |          |          |          |          |          |          |          |          |          |
|                                | (-1.464) | (-1.459) | (-1.357) |          |          |          |          |          |          |          |          |          |          |          |          |
| DUMMY$_{it}$                   | 0.0104** | 0.0103** |          |          |          |          |          |          |          |          |          |          |          |          |          |
|                                | (2.379)  | (2.350)  |          |          |          |          |          |          |          |          |          |          |          |          |          |
| DEM$_{it}$                     | -0.0109***| -0.0109***| -0.0109***| -0.0109***| -0.0105***|          |          | -0.0105***|          |          |          |          |          |          |          |
|                                | (-4.385) | (-4.370) | (-4.367) | (-4.367) | (-4.182) |          |          | (-4.182) |          |          |          |          |          |          |          |
| LIQA$_{it}$                    | -0.0022***| -0.0022***| -0.0022***| -0.0022***| -0.0022***|          |          | -0.0022***|          |          |          |          |          |          |          |
|                                | (-7.093) | (-7.068) | (-7.061) | (-7.061) | (-7.214) |          |          | (-7.214) |          |          |          |          |          |          |          |
| LIQH$_{it}$                    | 0.0002***| 0.0002***| 0.0002***| 0.0002***| 0.0002***|          |          | 0.0002***|          |          |          |          |          |          |          |
|                                | (4.759)  | (4.701)  | (4.699)  | (4.699)  | (4.569)  |          |          | (4.569)  |          |          |          |          |          |          |          |
| INF$_{it}$                     | -0.0001  | -0.0001  | -0.0001  | -0.0001  | -0.0001  |          |          | -0.0001  |          |          |          |          |          |          |          |
|                                | (-1.490) | (-1.515) | (-1.505) | (-1.505) | (-1.462) |          |          | (-1.462) |          |          |          |          |          |          |          |
| RISK$_{it}$                    | 0.0124***| 0.0124***| 0.0124***| 0.0124***| 0.0116***|          |          | 0.0116***|          |          |          |          |          |          |          |
|                                | (2.865)  | (2.777)  | (2.806)  | (2.806)  | (2.629)  |          |          | (2.629)  |          |          |          |          |          |          |          |
| EX$_{it}$                      | -0.1245***| -0.1223***| -0.1220***| -0.1220***| -0.1190***|          |          | -0.1190***|          |          |          |          |          |          |          |
|                                | (-4.391) | (-4.314) | (-4.305) | (-4.305) | (-4.199) |          |          | (-4.199) |          |          |          |          |          |          |          |
| $R^2$                          | 0.850    | 0.851    | 0.851    | 0.850    | 0.851    | 0.851    | 0.851    | 0.854    | 0.855    | 0.855    | 0.855    | 0.855    | 0.855    |          |

(continued)
Table 4. (continued)

| Variable | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     | (7)     | (8)     | (9)     | (10)    | (11)    | (12)    | (13)    | (14)    | (15)    |
|----------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Panel B: PLS method |         |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
| IVOL<sub>it</sub> | 0.5283*** | 0.4694*** | 0.4725*** | 0.4830*** | 0.4679*** | 0.4830*** | 0.5852*** | 0.4115**  | 0.3624**  | 0.3810*** | 0.4529*** | 0.4637*** | 0.3810*** | 0.4529*** | 0.4637*** | 0.3810*** |
| (3.179)  | (2.827) | (2.847) | (2.911) | (2.813) | (2.852) | (3.390) | (2.295) | (2.022) | (2.084) | (2.757) | (2.118) |         |         |         |         |
| Diff<sub>t−1</sub>S<sub>it−1</sub><sup>Tot</sup> | 0.0076*** | 0.0004 | 0.0075*** | 0.0063*** | 0.2471*** | (5.052) | (−0.909) | (4.995) | (4.239) |     |         |         |         |         |         |         |
| (3.179)  |         | (−0.909) |         |         | (2.070) |         |         |         |         |         |         |         |         |         |         |
| IVOL<sub>it</sub> × Diff<sub>t−1</sub>S<sub>it−1</sub><sup>Tot</sup> | 0.0075*** | 0.0019 | 0.0075*** | 0.0018 | 0.0074*** | 0.0066*** | 0.0072*** | −0.0019 | −0.0029 |     |         |         |         |         |         |         |
| (4.824)  | (−0.468) | (4.855) | (−0.442) | (4.823) | (3.932) | (4.615) | (−0.456) | (−0.721) |         |         |         |         |         |         |         |
| Diff<sub>t−1</sub>S<sub>it−1</sub><sup>Local</sup> | 0.2924** | 0.2906** | 0.2466 | 0.2451 | 0.2846** | 0.2769** | 0.2846** | 0.2769** |     |         |         |         |         |         |         |         |
| (−1.517) | (−1.493) | (−1.407) |         |         | (2.393) | (2.355) |         |         |         |         |         |         |         |         |         |
| s<sub>it−1</sub> | 0.0062 | 0.0052 | 0.0025 |     |         |         |         |         |         |         |         |         |         |         |         |
| (1.079)  | (0.901) | (0.433) |         |         |         |         |         |         |         |         |         |         |         |         |         |
| DUMMY<sub>it</sub> | 0.0106** | 0.0106** |     |         |         |         |         |         |         |         |         |         |         |         |         |
| (2.432)  | (2.423) |         |         |         |         |         |         |         |         |         |         |         |         |         |         |
| DEM<sub>it</sub> | −0.0109*** | −0.0105*** | −0.0105*** |     |         |         |         |         |         |         |         |         |         |         |         |
| (−4.385) | (−4.242) | (−4.240) |         |         |         |         |         |         |         |         |         |         |         |         |
| LIQA<sub>it</sub> | −0.0022*** | −0.0021*** | −0.0021*** |     |         |         |         |         |         |         |         |         |         |         |         |
| (−7.093) | (−6.849) | (−6.839) |         |         |         |         |         |         |         |         |         |         |         |         |
| LIQH<sub>it</sub> | 0.0002*** | 0.0002*** | 0.0002*** |     |         |         |         |         |         |         |         |         |         |         |         |
| (4.759)  | (4.748) | (4.735) |         |         |         |         |         |         |         |         |         |         |         |         |
| INF<sub>it</sub> | −0.0001 | −0.001* | −0.001 |     |         |         |         |         |         |         |         |         |         |         |         |
| (−1.490) | (−1.648) | (−1.625) |         |         |         |         |         |         |         |         |         |         |         |         |
| RISK<sub>it</sub> | 0.0124*** | 0.0112*** | 0.0114*** |     |         |         |         |         |         |         |         |         |         |         |         |
| (2.865)  | (2.569) | (2.627) |         |         |         |         |         |         |         |         |         |         |         |         |
| EX<sub>it</sub> | −0.1245*** | −0.1196*** | −0.1190*** |     |         |         |         |         |         |         |         |         |         |         |         |
| (−4.391) | (−4.225) | (−4.200) |         |         |         |         |         |         |         |         |         |         |         |         |
| R² | 0.850 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.851 | 0.855 |

Note. The stock price of A-shares over H-shares, expressed in RMB, is the dependent variable. Each company’s idiosyncratic risk is denoted by IVOL<sub>it</sub>. Global sentiment (S<sub>Tot−1</sub><sup>Global</sup>) is the principal component of the total sentiment indices (S<sub>Tot−1</sub>). Local sentiment (S<sub>Tot−1</sub><sup>Local</sup>) is the residual of the regression: S<sub>Tot−1</sub><sup>Local</sup> = α<sub>Tot−1</sub>' + β<sub>Tot−1</sub> S<sub>Tot−1</sub><sup>Total</sup>. Diff<sub>t−1</sub>S<sub>it−1</sub><sup>Tot</sup> = S<sub>t−1</sub><sup>Tot−1</sub> − S<sub>t−1</sub><sup>Total−1</sup>, Diff<sub>t−1</sub>S<sub>it−1</sub><sup>Total</sup> = S<sub>t−1</sub><sup>Total−1</sup> − S<sub>t−1</sub><sup>Total−1</sup> DEM<sub>it</sub>, LIQA<sub>it</sub>, LIQH<sub>it</sub>, INF<sub>it</sub>, RISK<sub>it</sub>, EX<sub>it</sub>, and DUMMY<sub>it</sub> are control variables defined above. S<sub>Tot−1</sub><sup>Total</sup> is BW index. The sample includes observations from January 2003 to December 2018. PCA = principal component analysis; PLS = partial least squares. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.
summarize, idiosyncratic risk is favorably correlated with the cross-listed company price premium. Furthermore, we find that a higher total (local) sentiment differential reinforces the sensitivity of idiosyncratic risk to the stock price premium in our study.

Conclusions

Although investor sentiment and idiosyncratic risk have been extensively studied, previous studies did not decompose the sentiment index into different fragments using both PCA and PLS approaches and combined it with idiosyncratic risk to investigate the price difference between A-shares and H-shares in China. In evaluating the price disparity, we compare the results of PCA and PLS investor sentiment indices. In our research, we discover that the price premium of the two types of shares is closely related to the total (local) investor sentiment differential. In addition, we discover that the idiosyncratic risk of the stock is positively linked to the price premium, and a higher sentiment differential strengthens the impact of the idiosyncratic risk of the stock on the price premium. Our article adds to the current literature on idiosyncratic risk and investor sentiment by decomposing the investor sentiment using PCA and PLS approaches and analyzing the interaction effect of idiosyncratic risk and sentiment in financial markets.

Findings in this article include much food for thought. First, in many ways, investor sentiment could be used as an indicator to help investors track market drift patterns. According to our findings, when investor sentiment differential rises, the price premium increases. We discover that the price premium will continue to climb for a few months when the sentiment differential hits its apex and starts to decrease. We assume that the price premium is 2 to 3 months ahead of the shift in sentiment differential. The lead–lag time intervals can be used by investors to make changes to their positions.

Second, Shanghai-Hong Kong Stock Connect has failed to narrow the price gap between A- and H-shares as predicted. On one hand, there exists a big difference between the two markets in terms of investor structures, regulatory environment, and investment philosophy. Different attitudes toward speculating and investing lead to different portfolio choices. Investors in Mainland China are most likely to focus on the stock price rather than the dividend paid and thereby lead to the pursuit of some particular stocks. The difference in the investment philosophy could not be changed with the launch of the program. On the other hand, the relatively strict restrictions result in limited investment choices for investors in Mainland China, which incurs long-term overvalued A-shares. Besides, the price of A-shares is more volatile and more susceptible to government policies. Overseas investors are prone to invest in A-shares to make profits. The above issues, in our view, would lead to a larger price gap between the two types of shares. It is worth noting that the program’s introduction seeks to promote the openness of the stock market, and narrowing down the price disparity is not the main target of the program. The diathesis of small and medium investors will be strengthened with the gradual opening of the capital market, and a more rational investment atmosphere will be formed under the guidance of the government. The Nasdaq-like Star Market, which debuted in July 2019, is a good example of this transition.

One of the drawbacks of this article is that it only includes companies listed in Shanghai, and hence only Shanghai-Hong Kong Stock Connect is added to our analysis. Since Shenzhen-Hong Kong Stock Connect has already launched on December 5, 2016, the discrepancies between the two programs could be explored in the following study. Furthermore, the influence of stock market transparency on price disparity, as well as investor sentiment and idiosyncratic risk, will be investigated in future studies.

Acknowledgments

We thank several participants in various seminars for their nice feedback and suggestions. All possible remaining errors are ours.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This project is funded by the research grant SM201910020003 of the Beijing Municipal Commission of Education. We gratefully acknowledge this support.

ORCID iD

Yuan Li https://orcid.org/0000-0002-7397-9032

References

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 5(1), 31–56.
Arquetti, G. C., Brown, W. O., & Burdekin, R. C. K. (2008). US ADR and Hong Kong H-share discounts of Shanghai-listed firms. Journal of Banking & Finance, 32(9), 1916–1927.
Bailey, W., & Jagtiani, J. (1994). Foreign ownership restrictions and stock prices in the Thai capital market. Journal of Financial Economics, 36(1), 57–87.
Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. Journal of Financial Markets, 7(3), 271–299.
Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. The Journal of Finance, 61(4), 1645–1680.
Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. Journal of Economic Perspectives, 21(2), 129–152.
Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. Journal of Financial Economics, 104(2), 272–287.
Beckmann, K. S., Ngo, T., & Wang, D. (2015). The informational content of ADR mispricing. *Journal of Multinational Financial Management, 32–33*, 1–14.

Ben-Rephael, A., Kandel, S., & Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics, 104(2)*, 363–382.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of Empirical Finance, 11*(1), 1–27.

Brown, G. W., & Cliff, M. T. (2005). Investor sentiment and asset valuation. *The Journal of Business, 78*(2), 405–440.

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), 325–356.

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

Chan, K., Menkveld, A. J., & Yang, Z. (2008). Information asymmetry and asset prices: Evidence from the China foreign share discount. *The Journal of Finance, 63*(1), 159–196.

Chi, L., Zhuang, X., & Song, D. (2012). Investor sentiment in the Chinese stock market: An empirical analysis. *Applied Economics Letters, 19*, 345–348.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARs investor sentiment and asset prices. *The Review of Financial Studies, 28*(1), 1–32.

Darrat, A. F., Elkal, K., & McCallum, B. (2006). Finance and macroeconomic performance. Some evidence for emerging markets. *Emerging Markets Finance and Trade, 42*(3), 5–28.

De Jong, A., Rosenthal, L., & Van Dijk, M. A. (2009). The risk and return of arbitrage in dual-listed companies. *Review of Finance, 13*(3), 495–520.

Fernald, J., & Rogers, J. (2002). Puzzles in the Chinese stock market. *The Review of Economics and Statistics, 84*, 416–432.

Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics, 88*(2), 299–322.

Froot, K. A., & Dabora, E. M. (1999). How are stock prices affected by the location of trade? *Journal of Financial Economics, 53*(2), 189–216.

Fu, J., Zhou, Q., Liu, Y., & Wu, X. (2020). Predicting stock market crises using daily stock market valuation and investor sentiment indicators. *The North American Journal of Economics and Finance, 51*, Article 100905.

Gagnon, L., & Karolyi, G. A. (2010). Multi-market trading and arbitrage. *Journal of Financial Economics, 97*(1), 53–80.

Grossmann, A., Ozuna, T., & Simpson, M. W. (2007). ADR mispricing: Do costly arbitrage and consumer sentiment explain the price deviation? *Journal of International Financial Markets, Institutions and Money, 17*(4), 361–371.

Guo, T. (2013). The Price discounts of Chinese cross-listed companies and their variation across sectors. *CMC Senior Theses Paper.*

Han, X., & Li, Y. (2017). Can investor sentiment be a momentum time-series predictor? Evidence from China. *Journal of Empirical Finance, 42*, 212–239.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *Review of Financial Studies, 28*(3), 791–837.

Karolyi, G. A., Li, L., & Liao, R. (2009). A (partial) resolution of the Chinese discount puzzle: The 2001 deregulation of the B-share market. *Journal of Financial Economic Policy, 1*(1), 80–106.

Kelly, B., & Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance, 68*(5), 1721–1756.

Kelly, B., & Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics, 186*(2), 294–316.

Khan, M. A., Hernandez, J. A., & Shahzad, S. J. H. (2020). Time and frequency relationship between household investors’ sentiment index and US industry stock returns. *Finance Research Letters, 36*, Article 101318.

Klemola, A. (2020). Internet search-based investor sentiment and value premium. *Finance Research Letters, 33*, Article 101224.

Kling, G., & Gao, L. (2008). Chinese institutional investors’ sentiment. *Journal of International Financial Markets, Institutions and Money, 18*(4), 374–387.

Kumar, A., & Lee, C. M. C. (2006). Retail investor sentiment and return comovements. *The Journal of Finance, 61*(5), 2451–2486.

Lan, Y., Huang, Y., & Yan, C. (2021). Investor sentiment and stock price: Empirical evidence from Chinese SEOs. *Economic Modelling, 94*, 703–714.

Lennon, M., & Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies, 19*(4), 1499–1529.

Li, Y., & Ran, J. (2020). Investor sentiment and stock price premium validation with Siamese twins from China. *Journal of Multinational Financial Management, 57–58*, Article 100655.

Mei, J., Scheinkman, J., & Xiong, W. (2005). Speculative trading and stock prices: Evidence from Chinese A-B share premia. *NBER, Working Paper.*

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance, 42*(3), 483–510.

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

Qiu, L., & Welch, I. (2004). Investor sentiment measures. *NBER, Working Paper.*

Ruan, Q., Wang, Z., Zhou, Y., & Lv, D. (2020). A new investor sentiment indicator (ISI) based on artificial intelligence: A powerful return predictor in China. *Economic Modelling, 88*, 47–58.

Rui, O., Wu, W., & Lee, B.-S. (2007). Market segmentation and stock prices discount in the Chinese stock market: Revisiting B-share discounts in the Chinese stock market. *Asia-Pacific Journal of Financial Studies, 37*, 1–40.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance, 16*(3), 394–408.

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance, 52*(1), 35–55.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance, 70*(5), 1903–1948.

Stulz, R. M. (1981). A model of international asset pricing. *Journal of Financial Economics, 9*(4), 383–406.

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. S. (2000). The effect of market segmentation on stock prices: The China syndrome. *Journal of Banking & Finance*, 24(12), 1875–1902.

Verma, R., & Soydemir, G. (2009). The impact of individual and institutional investor sentiment on the market price of risk. *The Quarterly Review of Economics and Finance*, 49(3), 1129–1145.

Wang, S. S., & Jiang, L. (2004). Location of trade, ownership restrictions, and market illiquidity: Examining Chinese A- and H-shares. *Journal of Banking & Finance*, 28(6), 1273–1297.

Wang, W. (2020). Institutional investor sentiment, beta, and stock returns. *Finance Research Letters*, 37, Article 101374.

Schwert, G. W. (2002). Stock volatility in the new millennium: how wacky is Nasdaq? *Journal of Monetary Economics*, 49(1), 3–26.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. In P. R. Krishnajah (Ed.), *Multivariate analysis* (pp. 391–420). Academic Press.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock (Ed.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). Seminar Press.

Wu, Q., Hao, Y., & Lu, J. (2017). Investor sentiment, idiosyncratic risk, and mispricing of American depository receipt. *Journal of International Financial Markets, Institutions and Money*, 51, 1–14.