Cross-listing A-Share and H-Share Stock Price Relationship Study: The Granger Causality Analysis Based on Establishment of Shanghai-Hong Kong Connect

by

Siyao Fang

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Business Honors Program

NYU Shanghai

May 2017

Professor Marti G. Subrahmanyam Professor Stephen Figlewski

Professor Jiawei Zhang

Faculty Advisers Thesis Adviser

Abstract

In this paper, we mainly focus on the informativeness and causality relationship detection based on the past historical data for those cross-listing companies have stocks both listed in Shanghai Stock Exchange and Hong Kong Stock Exchange. The paper applies Unit Root Test, Cointegration Test and Granger Causality Test to check whether long term stability exists between cross-listing companies' A and H shares. The results shows that At a 95% critical level, 25 companies reject null hypothesis of H share does not Granger Cause A share; 18 companies reject null hypothesis of A share does not Granger Cause H share. Among them, 5 companies show both informativeness from H share and A share to the other. The relationship may work as the first step to find possible trading strategy between A share and H share.

Acknowledgements

To Professor Stephen Figlewski – thank you for all your help from the very beginning of my project to the ending of this honor program. I can't have this project done without your kind suggestions on topic choosing, data process and test calibration. I'm truly appreciated the opportunity you provide and the time you spent with my callow research.

To Professor Marti Subrahmanyam – thank you for the whole arrangement of NYU Shanghai Business Honor Program. This is a wonderful two-semester course which brings a lot of fresh idea and novel thought to me.

To Professor Jiawei Zhang – thank you for kind suggestions when I am down in my first semester. Your guidance of my process points a clear way for me to further my honor program and catch up with my original plan.

1. Thesis Introduction:

Mainland companies have a long history to list stocks in both Shanghai Stock Exchange (hereafter SSE) and Hong Kong Stock Exchange (HKSE). According to perceptual and traditional financial view, same assets should be priced as same seems to be an obvious result for the cross-listing stocks. However, pricing spreads have also a long history starting from the very first day of dual listing. Is the empirical result wrong or something behind influences the "obvious" result? In this paper, we will show the history of pricing spread for cross-listing stocks in SSE and HKSE and try to using statistical methodology mainly cointegration test [22] and Granger Causality test [21] to explore the deeper relationship underlying the cross-listing stock price. This study hopes to work as a signal research to see that if any informativeness and causality relationship between those cross-listing companies' stock price in SSE and HKSE. And to see if any sign can show that the price spread will be more steady or fluctuant in future considering the establishment of Shanghai-Hong Kong Stock Connect. The result of the paper suggests that there is clearly informativeness and causality relationship between SSE and HKSE for those cross-listing stocks. Furthermore, according to the causality result categorization and industry investigation, the direction of information flow may have huge relationship with the different closing time between SSE and HKSE together with the industry of selected stocks.

2. Background:

Since 15th July, 1993, two years after the establishment of Shanghai Composite Index, Tsingtao Beer as the first mainland company listed in HKSE got public listing in Hang Seng Index, multiple mainland Chinese companies choose to issue their stock overseas to foreign

investors sharing the financial achievement from the reform and opening-up policy. HKSE advantaged from its policy and location, have become the most prior oversea exchange for mainland Chinese companies to considerate. After the first bell ringing of Hong Kong listed mainland company, tens of gigantic companies came to Hong Kong and issued their stock in Hong Kong Stock Exchange as a financing supplement after the public listing in Shanghai Stock Exchange.

There are 62 companies dual lists their companies' stock in both SSE and HKSE till now. Most of companies announce that their will consider stock listed in both market totally same. That is, for each market's stock, stock owners own same shares of the company, share same earning rights and voting right with their stocks. Dividends and shareholders' power are equal to investors from different market as well. (China Unicom so far is the only exception to the above stock regulation. China Unicom's H share stock is issued by buying an existing H share company. Therefore, its stocks represents different company in brief.) However, due to historical policy difference and strict limitation on investment, the price of same stock in two markets shows huge distinctions over decades. It is generally acknowledged that the price spreads of cross-listing stocks in SSE and HKSE are caused for the following reasons:

- Different Investor's Expectation and Risk Premium: H shares' investors are mostly foreigners while A shares' investors are almost Chinese investors. These different investors share different future expectation and different risk aversion towards market.
- Different Trading Regulation: in A share market, stocks are trading under T+1 rule, that is, any bought stock can't be sold until tomorrow. However, in H shares market, stock is tradable in anytime. Also, in A share market, a maximum 10% fluctuation is

established by China Securities Regulatory Commission.

- Different Government Supervision and Policy. A share are under the supervision of China Securities Regulatory Commission while H share are supervised by Hong Kong Securities and Futures Commission. Different governors are response to different policy makers who assign different policy based on different environment of Shanghai and Hong Kong.
- Different Currency: currency exchange is under strict control. RMB is the only official currency in SSE while HKD is the only official currency in HKSE. The limited exchange amount largely limited the exchange of stocks.
- Not cross-tradable: though the cross-listing stocks are same stocks, they are not cross-tradable in other market. That is, the stocks which bought from SSE as A shares, are not tradable in HKSE as H shares.

Based on those previous reasons, considering that the "market stock portfolio" that a Chinese investor is able to hold and the "world market portfolio" that a foreign investor can hold, H share provides the opportunities for a foreign investor to explore the Chinese stock market. This should usually makes H share cost a little higher than A share for same company's stock since it has substantial value for diversification for foreign investors. Later related work about American cross-listing stock also proved this [3]. While for a Chinese investor who is not able to own foreign stocks, A shares creates a lot of systematic nondiversifiable risk for them. With different investor populations, A and H shares also have different risk characteristics which makes daily volume and transaction price totally different between them.

Among these reasons, the difference investors' expectation, and the difference of currency

are expected to get integration thanks to the establishment of Shanghai-Hong Kong stock Connect on 17th Nov, 2014. The Connect enables investors to trade specified H shares in SSE and trade some SSE's A-shares in HKSE. Therefore, the difference of expectation from different investors is considered to be narrowed down. Furthermore, the liquidity problem of not cross-tradable stocks is largely solved by the channel provided by the Connect. Though, in real practice, the H shares are still not tradable in A share market, however, the establishment of the Connect has hugely improved the situation that H share and A share are completely in different market segment. It's reasonable to believe that the appearance of the Connect is a cross-market tool to further stabilize the market and reveal the true value of stocks under the globalized investors.

3. Related Work

Lots of work related to the cross-listing stock price have been published. In this section, we will provide you the closest previous work to out topic. The work will be generally separated into three part: Cross-listing Stock Relationship, Existing Arbitrage Study and Basic Statistical Fundaments.

- Cross-listing Stock Relationship. An enormous literatures have investigated cross-listing stock price relationship, however, most past investigations focus on stock market in the United States [1]. These investigations combine American market (New York Stock Exchange) to other well-known stock markets such as Hong Kong [2], Canadian [3], Germany [4], South Korean [5] and so on [6][7][8][9]. Most researchers find that American markets seems to be more informative than other market [3][5] which provide better adjustment to the American markets. Because of this, usually

foreign market pays a higher price to cross-listing stocks together with their better diversity to foreign investors. From this sense, A shares show completely different from American-related cross-listing stocks since A shares are usually more expensive to H shares after proper currency changes. A shares and H shares study has also appeared in recent years. However, not only the amount but also the conclusions are hard to apply in realistic practice. Study conducted by Chen [1], Cai et al. [10][12], Su et al. [11] all applied cointegration as the method to study relationship between A shares and H shares. However, most of them only focus on aggregation index (Shanghai Composite Index and Hang Seng Index) which has little practical guidance in real trading strategy. In these past studies, most of them apply the Engle-Granger two-step cointegration method [23] and Granger Causality test [21]. In this paper, we will also apply those method for further investigation.

- Existing Arbitrage Study. Arbitrage from financial derivatives has been studied since early ages and multiple mature strategies have been widely accepted in real trading activities such as put-call parity. Unlike financial derivatives, which has risk-free arbitrage strategy, the existing of market segment [15] prevents the risk-free arbitrage due to the limit of arbitrage [16]. According to Gordon Growth Model [17], expectation and discount rate influence stock price with the most prior. Discount rate, as Black et al. suggested in classical Capital Asset Pricing Model [18], is highly related to risk-free rate and risk premium, which are both hard to unify in different market. Modern researchers have therefore, mainly focused on the causality and informativeness study, expecting that the price movement in one market can be as a

guidance for the price movement in the other [19][20]. Further studies also propose possible arbitrage strategy based on data analysis on cross-listing stocks [13][14]. In A-share and H-share case, two methods are widely considered as arbitrage opportunities: 1. By retardance of closing time. HKSE close at one hour later than SSE. Thus, the closing price in SSE may have some guidance to the closing price in HKSE. Also, in another hand, the closing price in HKSE may have some guidance to the opening price in SSE in next trading day. If the latter is true, the lag information should be small. 2. By stock price relationship. This is what most statistical researches try to reveal.

- Basic Statistical Fundaments. Cointegration is a statistical property for given time series if there exists a linear combination which is integrated of order zero. The property of cointegration is first mentioned and analyzed by Engle et al. [23] and feasible solution to the calculation of integration is also recorded in the same paper. Later researchers, including Johansen [14] and Phillips et al. [15] both provide their solution method to the calculation of cointegration, especially to multiple time series and to high order of integration. Current cointegration applications and researches have been widely accepted in every time series analysis field, such as social science [25], finance [26] and so on. The reliability and stability of cointegration has been long researched and proved.

  Granger Causality test is a statistical hypothesis test to check whether given two time series data have certain determined causality relationship found by Granger in 1969 [21]. Causality relationship between given time series is based on two aspects

according to the Granger Causality test: 1. The cause happens with a certain lag variable, that is, the cause will happen before the effect. 2. Each cause will have unique information to the effect on the predicted value. With these two hypothesis, the test will show the t-statistical result of the null hypothesis so that the causality relationship can be checked with possibility table.

4. Data Summary and Process

The daily price dataset we gathered is come from iFind database of Tonghuashun Ltd. with 59 cross-listing companies, exclude China Unicom and COSCO Corporation. The companies cover over 10 industries from finance to manufacture, from resources to food. The daily RMB to HKD exchange rate is gathered from Xignite database of Xignite. Data cleaning involves: 1. Convert all H shares' closing price from HKD to RMB. 2. Delete all trading days that fit any of following regulations: Either Shanghai or Hong Kong market is closed; Isolated trading days which are not suitable for lag investigation. In order to find the relationship within a reasonable time period, the trading days we used are from Jun 1st, 2012 to Jun 31st, 2016, 61 months in total. The number of maximum legitimate trading days during this period is 966 trading days that 20 companies achieve this number. The number of minimum legitimate trading days is 590 trading days that happens only to CEB bank, a newly issued commercial bank.

Considering the huge dataset, in this paper, we will apply Ping An of China as the example for every following statistical steps since Ping An of China shows a clearly correlation and the results of each tests is well organized to work as an example.

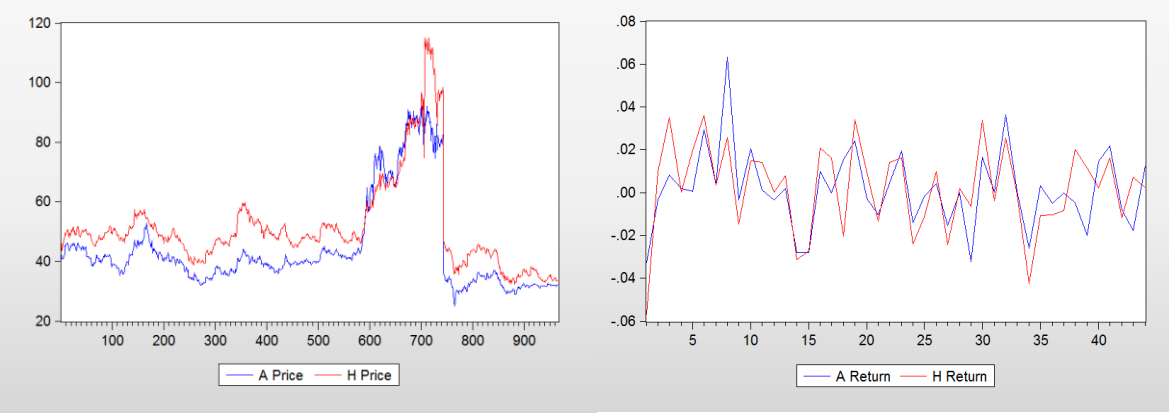


Fig.1 A, shows the daily closing price of Ping An of China in both A share (red) and H share (blue). B, shows the logged daily return of Ping An of China in both A share (red) and H share (blue) in the first 45 trading days. Both figures show that they are highly relevant to each other, same trend in short run and clear synchronization on price movement.

From Fig.1, A, shows the daily closing price of Ping An of China in both A share (red) and H share (blue). B, shows the logged daily return of Ping An of China in both A share (red) and H share (blue) in the first 45 trading days because the total return movement chart is messy and hard to distinguish the relationship with crowded data moving randomly up/down to 0. In this research the actual data we analysis is the logged daily return in order to eliminate nonlinear influence. Both figures show that they are highly relevant to each other, same trend in short run and clear synchronization on price movement. In chart a, we can also see clear share split in both A share and H share, same split rate also proves that both share are identically the same. The basic statistical results shows that there are 965 logged return for Ping An of China with a correlation of 0.776775 between A share's logged return and H share's logged return, which is very good among all companies that the paper has analyzed. There are some trading days have logged return over 10% or less than -10%. The excessive high one is caused by the elimination of unqualified trading days and the excessive low one is caused by the share split.

We apply basic least squared regression model:

$$\text{A return}_t = \beta \text{H return}_t + C + \varepsilon$$

to check the basic relationship with result as following:

Dependent Variable: A_RETURN03
Method: Least Squares
Date: 04/12/17 Time: 23:02
Sample (adjusted): 1 965
Included observations: 965 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
H_RETURN03	0.786758	0.020555	38.27552	0.0000
C	-2.34E-05	0.000699	-0.033452	0.9733

R-squared	0.603380	Mean dependent var	-0.000281
Adjusted R-squared	0.602968	S.D. dependent var	0.034471
S.E. of regression	0.021720	Akaike info criterion	-4.819075
Sum squared resid	0.454313	Schwarz criterion	-4.808977
Log likelihood	2327.204	Hannan-Quinn criter.	-4.815231
F-statistic	1465.016	Durbin-Watson stat	2.340527
Prob(F-statistic)	0.000000		

Fig.2 The Least Squared Regression Result: The t-statistical result of coefficient of H return rejects the null hypothesis that the coefficient of H return is 0. The constant variable C gets a result close to 0 is also proved by t-statistical test.

Date: 04/12/17 Time: 23:16
Sample: 1 966
Included observations: 965

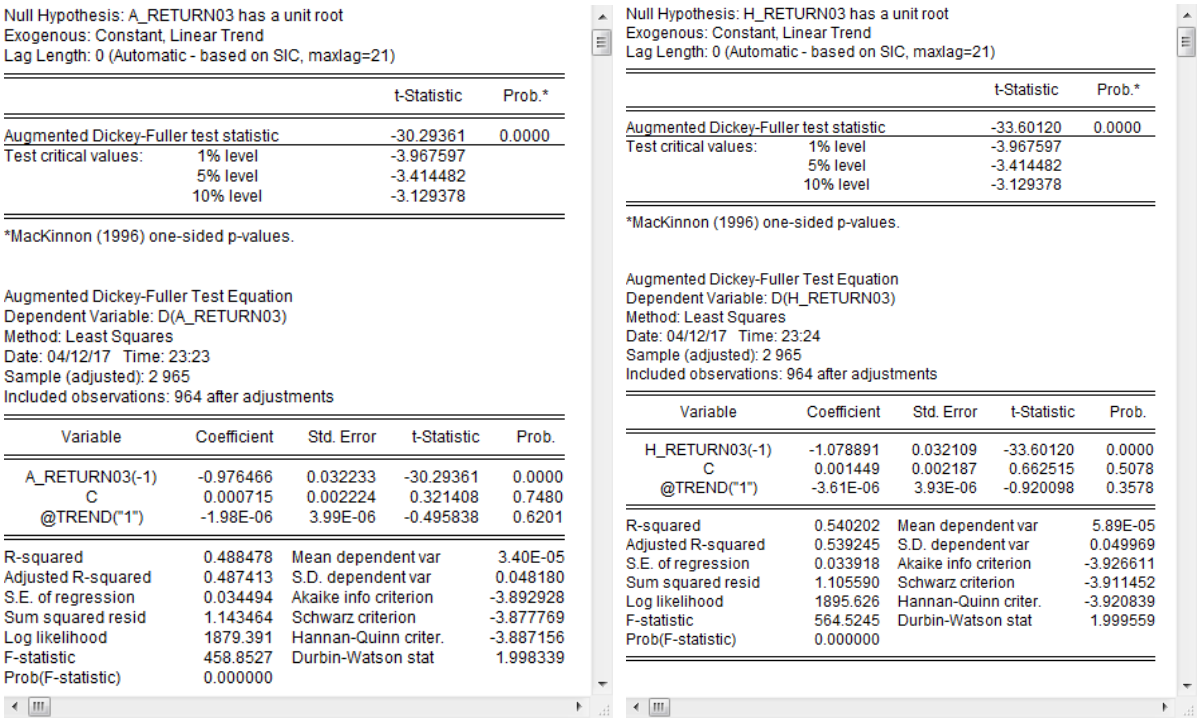
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.170	-0.170	28.114	0.000
		2	0.015	-0.015	28.326	0.000
		3	-0.059	-0.061	31.724	0.000
		4	0.014	-0.007	31.906	0.000
		5	-0.052	-0.053	34.506	0.000

Fig.3 The Autocorrelation Result: No clear autocorrelation is found with a maximum lag of 5

The t-statistical result of coefficient of H return rejects the null hypothesis that the coefficient of H return is 0. Based on the meaning of using lagged as variable, the coefficient of 0.78 shows that any 1% change in H shares will cause a 0.78% change in A shares. The constant variable C gets a result close to 0 is also proved by t-statistical test. The R-squared, 0.603, suggests that the fitness of this basic regression is not bad especially after the comparison with other companies. In order to prevent autocorrelation, we also check the autocorrelation of the residual from the regression model. Fig 3 shows the result of autocorrelation of residual with a maximum lag of 5. The lag is chosen based on: 1. The closing time difference suggests that the lag should not be big. 2. Five is a reasonable guess since a week has five trading days. The result of autocorrelation finds no clear autocorrelation within this regression. The regression may be reasonable and reliable.

In order to prevent spurious regression from two non-stationary series where the hidden

variable in two non-stationary series may cause the regression well, we also apply the unit root test to the data. The test also work as a preparation test for upcoming cointegration test.



Null Hypothesis: A_RETURN03 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=21)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-30.29361	0.0000
Test critical values:	1% level	-3.967597	
	5% level	-3.414482	
	10% level	-3.129378	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(A_RETURN03)
Method: Least Squares
Date: 04/12/17 Time: 23:23
Sample (adjusted): 2 965
Included observations: 964 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
A_RETURN03(-1)	-0.976466	0.032233	-30.29361	0.0000
C	0.000715	0.002224	0.321408	0.7480
@TREND("1")	-1.98E-06	3.99E-06	-0.495838	0.6201

R-squared	0.488478	Mean dependent var	3.40E-05
Adjusted R-squared	0.487413	S.D. dependent var	0.048180
S.E. of regression	0.034494	Akaike info criterion	-3.892928
Sum squared resid	1.143464	Schwarz criterion	-3.877769
Log likelihood	1879.391	Hannan-Quinn criter.	-3.887156
F-statistic	458.8527	Durbin-Watson stat	1.998339
Prob(F-statistic)	0.000000		

Null Hypothesis: H_RETURN03 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=21)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-33.60120	0.0000
Test critical values:	1% level	-3.967597	
	5% level	-3.414482	
	10% level	-3.129378	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(H_RETURN03)
Method: Least Squares
Date: 04/12/17 Time: 23:24
Sample (adjusted): 2 965
Included observations: 964 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
H_RETURN03(-1)	-1.078891	0.032109	-33.60120	0.0000
C	0.001449	0.002187	0.662515	0.5078
@TREND("1")	-3.61E-06	3.93E-06	-0.920098	0.3578

R-squared	0.540202	Mean dependent var	5.89E-05
Adjusted R-squared	0.539245	S.D. dependent var	0.049969
S.E. of regression	0.033918	Akaike info criterion	-3.926611
Sum squared resid	1.105590	Schwarz criterion	-3.911452
Log likelihood	1895.626	Hannan-Quinn criter.	-3.920839
F-statistic	564.5245	Durbin-Watson stat	1.999559
Prob(F-statistic)	0.000000		

Fig.4 A. shows the unit root test result of A return. The ADF test result rejects the null hypothesis that A return series has a unit root. The test result on trend and non-zero constant also reject the existence of trend and non-zero constant. B. shows the unit root test result of H return. The ADF test result rejects the null hypothesis that H return series has a unit root. The test result on trend and non-zero constant also reject the existence of trend and non-zero constant.

Fig.4 shows the unit root test result for both A return and H return, The ADF test result rejects the null hypothesis that A return series has a unit root. The test result on trend and non-zero constant also reject the existence of trend and non-zero constant. Also the ADF test result rejects the null hypothesis that H return series has a unit root. The test result on trend and non-zero constant also reject the existence of trend and non-zero constant. Both time series are I(0) stationary series that fit cointegration test.

In cointegration test, we apply the Engle-Granger two-step method, checking the unit root from the residuals of the least squared regression. Two is considering as the maximum lag since we believe that: first, the time influence is considering to be either on same day or

one day later; second, according to traditional empirical analysis, two is wildly accepted as the normal lag.

Null Hypothesis: RESID01 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-24.03999	0.0000
Test critical values:	1% level	-3.967607	
	5% level	-3.414487	
	10% level	-3.129381	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID01)
Method: Least Squares
Date: 04/13/17 Time: 11:28
Sample (adjusted): 3 965
Included observations: 963 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID01(-1)	-1.187439	0.049394	-24.03999	0.0000
D(RESID01(-1))	0.014684	0.032281	0.454884	0.6493
C	-0.000332	0.001385	-0.239482	0.8108
@TREND("1")	6.81E-07	2.49E-06	0.273819	0.7843

R-squared	0.585280	Mean dependent var	1.16E-05
Adjusted R-squared	0.583982	S.D. dependent var	0.033238
S.E. of regression	0.021438	Akaike info criterion	-4.843118
Sum squared resid	0.440762	Schwarz criterion	-4.822890
Log likelihood	2335.961	Hannan-Quinn criter.	-4.835416
F-statistic	451.1341	Durbin-Watson stat	2.001625
Prob(F-statistic)	0.000000		

Null Hypothesis: RESID01 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Fixed)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-20.17057	0.0000
Test critical values:	1% level	-3.967617	
	5% level	-3.414492	
	10% level	-3.129383	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID01)
Method: Least Squares
Date: 04/13/17 Time: 11:28
Sample (adjusted): 4 965
Included observations: 962 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID01(-1)	-1.259662	0.062450	-20.17057	0.0000
D(RESID01(-1))	0.085594	0.049701	1.722180	0.0854
D(RESID01(-2))	0.060869	0.032243	1.887823	0.0594
C	-0.000263	0.001386	-0.189518	0.8497
@TREND("1")	5.81E-07	2.49E-06	0.233892	0.8151

R-squared	0.587207	Mean dependent var	2.04E-05
Adjusted R-squared	0.585481	S.D. dependent var	0.033254
S.E. of regression	0.021410	Akaike info criterion	-4.844716
Sum squared resid	0.438685	Schwarz criterion	-4.819409
Log likelihood	2335.308	Hannan-Quinn criter.	-4.835079
F-statistic	340.3377	Durbin-Watson stat	2.000933
Prob(F-statistic)	0.000000		

Fig.5 the t-statistic result rejects the null hypothesis that the residual has a unit root for neither lag 1 nor lag 2. Here the residual is I(0) stationary series. We choose lag as 2 according to their AIC (Akaike information criterion) value

From Fig.5, the t-statistic rejects the null hypothesis that the residual has a unit root for neither lag 1 nor lag 2. Here the residual is I(0) stationary series. We choose lag as 2 according to their AIC (Akaike information criterion) value [27]. The later Granger Causality test will follow the lag we get here.

At last, we do the Granger Causality test to the dataset, the result for two null hypothesis is as following:

H return does not Granger Cause A return: Prob. 0.2339

A return does not Granger Cause H return: Prob. 1.E-07

The F-statistic clearly rejects the null hypothesis that A shares doesn't Granger Cause H shares. However, the same test can't reject the null hypothesis that H share doesn't Granger

Cause A shares. Considering the above discussion about different closing time of A share and H share, this may be a good answer and evidence to the price adjustment from A share closing price.

5. Result Summary

After applying all steps in above to all 59 companies, the Granger Causality result are showing in the following table:

	P-value that H_RETURN does not Granger Cause A_RETURN	P-value that A_RETURN does not Granger Cause H_RETURN
华能国际 HPI	0.007	0.1171
招商银行 CMB	0.0309	0.0000007
工商银行 ICBC	0.2695	0.0000005
中国平安 PING AN OF CHINA	0.2339	0.0000001
中国神华 China Shenhua	0.02	0.0011
交通银行 BANKCOMM	0.0076	0.149
农业银行 AGRICULTURAL BANK OF CHINA	0.18	0.0449
长城汽车 GREAT WALL MOTOR	0.9289	0.4412
中国银行 BANK OF CHINA	0.0892	0.0497
上海医药 Shanghai Pharma	0.1661	0.1259
民生银行 CMBC	0.3393	0.5107
建设银行 CCB	0.0003	0.05
华电国际 HDPI	0.0146	0.5233
中国人寿 CHINA LIFE	0.7136	0.4365
光大银行 CEB BANK	0.16	0.9982
白云山 GYBYS	0.0002	0.5354
新华保险 NCI	0.8313	0.4637
中信证券 CITIC Securities Co.,Ltd.	0.0021	0.6522
海通证券 Haitong Securities	0.0004	0.1958
广深铁路 GSRC	0.188	0.00000002
广汽集团 GSRC	0.7758	0.6912
中信银行 CNCB	0.9022	0.04
中国国航 AIR CHINA LTD.	0.2245	0.7875
中国石油 PETROCHINA	0.3327	0.0681
金隅股份 BBGM	0.9293	0.0011
中国交建 CCCC	0.6466	0.2669

南方航空 CSN	0.2595	0.0088
江西铜业 JCCL	0.8423	0.0012
东方电气 DEC	0.0131	0.5594
东方航空 CEA	0.0192	0.2375
中国中铁 CHINA RAILWAY	0.0083	0.0052
中煤能源 CHINA COAL ENERGY	0.6753	0.0487
中船防务 COMEC	0.0197	0.6746
郑煤机 ZMJ	0.0018	0.0695
中国铁建 CRCC	0.0244	0.0933
大唐发电 DATANG POWER	0.0106	0.5512
中海油服 COSL	0.1119	0.6008
紫金矿业 ZIJING MINING	0.0254	0.4074
中国铝业 CHALCO	0.0369	0.574
四川成渝 SICHUAN EXPRESS	0.1246	0.5026
马钢股份 MAS C.L.	0.4362	0.9893
深高速 SZEW	0.2192	0.0879
大连港 DALIAN PORT	0.1535	0.4902
南京熊猫 NPEC	0.0389	0.3659
北辰实业 BEIJING NORTH STAR	0.0019	0.2748
中国中冶 MCC	0.6056	0.7092
上海石化 SPC	0.0135	0.0935
重庆钢铁 CISL	0.0095	0.6864
上海电气 SHANGHAI ELECTRIC	0.0002	0.0005
洛阳钼业 CMOC	0.6997	0.3665
昆明机床 KMTCL	0.1703	0.0669
兖州煤业 YANZHOU COAL	0.0535	0.0465
创业环保 TCEPC	0.377	0.3379
洛阳玻璃 OGI	0.00006	0.0871
一拖股份 FRIST TRACTOR	0.3801	0.6998
海螺水泥 ACC	0.1823	0.02
青岛啤酒 TWINGTAO BREWERY	0.0107	0.0004
中国太保 CPIC	0.1385	0.6273
宁沪高速 JIANGSU EXPRESS	0.3476	0.0007

From the Granger Causality test, we find 25 companies' results reject the null hypothesis that H share does not Granger Cause A share on 5% level with their chosen lag and 12 companies can reject the null hypothesis on 1% level that indicate a strong evidence that the H share price may have a long term influence on A share price.

Meanwhile 18 companies' results reject the null hypothesis that A share does not Granger

Cause H share on 5% level with their chosen lag, Also, 12 companies can reject the null hypothesis on 1% level which indicate a strong evidence that the A share price may have a long term influence on H share price.

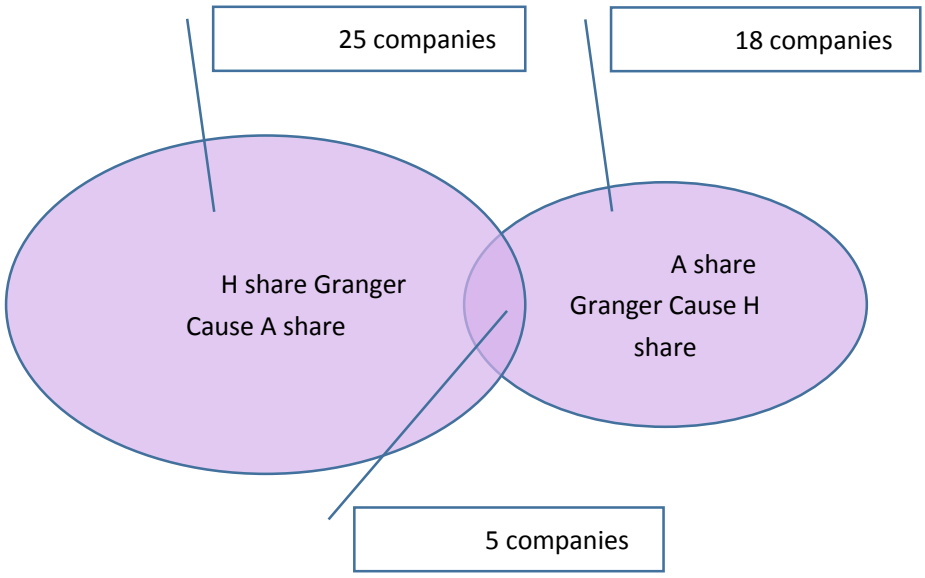


Fig. 6: Venn diagram of the final results from all 59 companies. At a 95% critical level, 25 companies reject null hypothesis of H share does not Granger Cause A share; 18 companies reject null hypothesis of A share does not Granger Cause H share. Among them, 5 companies show both informativeness from H share and A share to the other.

Here, we use a Venn diagram to illustrate the final results and its composition in Fig. 6. From the above Venn diagram at 5% critical point, we can see that 25 companies indicate that H share may have long term impact on A share. Interestingly, after searching the industry categories, we find that all security companies, Citic Securites and Haitong Securities show in this category. A reasonable guess is that securities companies will be influenced more by the oversea information, especially information after the closing time of A share market. Also, most resources companies show up in this category, such as CISL, SPC and Zijing Mining. Natural resources are also considering as important part of the globalization trading system that may influence the price a lot. Correspondently, 18 companies indicate that A share may have long term impact on H share. We also exam the industry distribution of those

companies. Almost all bank companies show in this category. A reasonable explanation may be that Chinese bank will be influenced more by mainland information since all Chinese commercial bank is under the governance of People's Bank of China which is under the governance of Chinese Government. There are also five companies show both long term informative relationship from A share and H share. They are CMB, China Shenhua, Twingtao Brewery, China Railway and Shanghai Electric. The reason of these five companies are not sure till now since they are in different industries with different history in cross-listing market.

Overall, the results show two main conclusions towards the informativeness and causality relationship between cross-listing stock in both A share market and H share market. First, the informativeness and causality do exist between cross-listing companies. 38 out of 59 companies show strong Granger Cause between two markets. Also, the lag factor for each company between H share and A share shows that the previous guess about closing time influence do hugely impact each stock price.

Based on what we've found, with the help of the Connect, investors from both market will have larger chance to make profit in the following four situation in which situation II and IV is not trading practicable before the connect:

I. A Chinese investor when A shares are higher than H shares for a given company. The Chinese investor will try to sell A share stock and buy H share if one shares have certain informative power to the other shares price. This kind of cash flow will calm down the A share price together with raise the H share price which will help to equalize both prices.

II. A Chinese investor when A shares are lower than H shares.

The Chinese investor will try to buy A shares and sell H shares simultaneously to make profit if one shares have certain informative power to the other shares price.

III. A foreign investor when A shares are higher than H shares for a given company.

Same as Chinese investors in situation I.

IV. A foreign investor when A shares are lower than H shares.

Same as Chinese investors in situation II.

6. Future Work and Expectation

According to our two assumption about closing time impact of A share and H share, it is reasonable for us to check the relationship between H share's closing price with the A share's opening price in order to see if any further data can show that the informativeness flows from H share's closing price to A share's opening price. Secondly, the study need more work to come up with a long term regression model according to its Granger Causality test result. The regression model may not have guidance in real arbitrage practice trading, however, the regression model will largely help to explore more about the inner relationship.

We also examine the stock performance before and after the Connect is established, however, the result doesn't show too much different from what we get through the above steps. One possible reason may be that the market efficiency is considerably enough even include the data before the Connect.

In Aug, 2016, China Security Committee announced that they would abolish the 13 billion RMB daily trading limitation. This regulation will further low the limitation of stock exchange and help to explore real value to those cross-listing companies' stocks. In long

term, the relationship between A share and H share might be much steady with a more open and efficiency market.

Citation

[1] Haiqiang Chen & Yanli Zhu (2015) An empirical study on the threshold cointegration of Chinese A and H cross-listed shares, *Journal of Applied Statistics*, 42:11, 2406-2419, DOI: 10.1080/02664763.2015.1034660

[2] Xu, X. E., & Fung, H. G. (2002). Information flows across markets: evidence from China–backed stocks dual–listed in Hong Kong and New York. *Financial review*, *37*(4), 563-588.

[3] C.S. Eun, and S. Sabherwal, *Cross-border listings and price discovery: Evidence from US-listed Canadian stocks*, J. Finance 58(2) (2003), pp. 549–576.

[4] J. Grammig, M. Melvin, and C. Schlag, *Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects*, J. Empir. Finance 12(1) (2005), pp. 139–164.

[5] K.L. Lee. *A cointegration analysis of price transmission between ADRs and dually listed South Korean stocks*, Math. Comput. Simulation 78 (2008), pp. 200–208.

[6] Gompers, P. A., Ishii, J., & Metrick, A. (2004). *Incentives vs. control: An analysis of US dual-class companies* (No. w10240). National Bureau of Economic Research.

[7] Licht, A. N. (2003). Cross-listing and corporate governance: bonding or avoiding. *Chi. J. Int'l L.*, *4*, 141.

[8] Lang, M. H., Lins, K. V., & Miller, D. P. (2003). ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value?. *Journal of Accounting Research*, *41*(2), 317-345.

[9] Foerster, S. R., & Karolyi, G. A. (1999). The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stocks listing in the United States. *The Journal of Finance*, *54*(3), 981-1013.

[10] C.X. Cai, P.B. McGuinness, and Q. Zhang, *The pricing dynamics of cross-listed securities: The case of Chinese A and H-shares*, J. Bank Finance. 35 (2011), pp. 2123–2136.

[11] Q. Su, T.T.L. Chong, and I.K.M. Yan, *On the convergence of the Chinese and Hong Kong stock markets: A cointegration analysis of the A and H shares*, Appl. Finance Econom. 17 (2007), pp. 1349–1357.

[12] 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.

[13] Blouin, J., Hail, L., & Yetman, M. H. (2009). Capital gains taxes, pricing spreads, and arbitrage: Evidence from cross-listed firms in the US. *The Accounting Review*, *84*(5), 1321-1361.

[14] Board, J., & Sutcliffe, C. (1996). The dual listing of stock index futures: arbitrage, spread arbitrage, and currency risk. *Journal of Futures Markets*, *16*(1), 29-54.

[15] Cano, C., "The Recent Evolution of Market Segmentation Concepts and Thoughts Primarily by Marketing Academics," in E. Shaw (ed) *The Romance of Marketing History*, Proceedings of the 11th Conference on Historical Analysis and Research in Marketing (CHARM), Boca Ranton, FL, AHRIM, 2003.

[16]Shleifer, Andrei; Vishny, Robert (1997). *"The limits of arbitrage". Journal of Finance. **52**: 35–55. CiteSeerX 10.1.1.184.9959. doi:10.1111/j.1540-6261.1997.tb03807.x.*
[17] Gordon, M.J and Eli Shapiro (1956) "Capital Equipment Analysis: The Required Rate of Profit," Management Science, 3,(1) (October 1956) 102-110. Reprinted in *Management of Corporate Capital*, Glencoe, Ill.: Free Press of, 1959.
[18] Black, Fischer., Michael C. Jensen, and Myron Scholes (1972). *The Capital Asset Pricing Model: Some Empirical Tests*, pp. 79–121 in M. Jensen ed., Studies in the Theory of Capital Markets. New York: Praeger Publishers.
[19] Li, S., Brockman, P., & Zurbruegg, R. (2015). Cross-listing, firm-specific information, and corporate governance: Evidence from Chinese A-shares and H-shares. *Journal of Corporate Finance*, *32*, 347-362.
[20] Kot, H. W., & Tam, L. H. (2016). Are stock price more informative after dual-listing in emerging markets? Evidence from Hong Kong-listed Chinese companies. *Pacific-Basin Finance Journal*, *36*, 31-45.
[21] Granger, C. W. J. (1969). "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". *Econometrica*. 37 (3): 424–438. doi:10.2307/1912791. JSTOR 1912791
[22] Nelson, C. R.; Plosser, C. R. (1982). "Trends and random walks in macroeconmic time series". *Journal of Monetary Economics.* 10 (2): 139. doi:10.1016/0304-3932(82)90012-5.
[23] Engle, Robert F.; Granger, Clive W. J. (1987). "Co-integration and error correction: Representation, estimation and testing". *Econometrica*. 55 (2): 251–276. JSTOR 1913236.
[24] Johansen, Søren (1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models". *Econometrica*. 59 (6): 1551–1580. JSTOR 2938278
Phillips, P. C. B.; Ouliaris, S. (1990). "Asymptotic Properties of Residual Based Tests for Cointegration". *Econometrica*. 58 (1): 165–193. JSTOR 2938339
[25] GOCER, I., & TATLI, H. (2017). Sustainability of social security systems in EU countries: Panel cointegration analysis with multiple structural breaks under cross-sectional dependence. *Theoretical & Applied Economics*, 24(1), 83-98.
[26] Siddiqui, T. A., & Rawal, I. (2017). Impact of Natural Disasters on Cointegration and Diversification: A Study of World's Major Stock Indices. *IUP Journal Of Applied Economics*, 6(1), 7-33.
[27] Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N.; Csáki, F., *2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971*, Budapest: Akadémiai Kiadó, pp. 267–281.