Doctoral Dissertation:

Analysis on dually listed companies in Hong Kong and China Stock Markets

Written by: Olivia Tze Kar Choi

THE INTERNATIONAL SCHOOL OF MANAGEMENT

PARIS • NEW YORK • SHANGHAI • NEW DELHI • SÃO PAULO • CAPE TOWN

DETAILS

Author Name: Olivia Tze Kar Choi, PhD

Year: 2019

Title: Analysis on dually listed companies in Hong Kong and China Stock Markets

Document type: Doctoral dissertation

Institution: The International School of Management (ISM)

URL: https://ism.edu/images/ismdocs/dissertations/choi-phd-dissertation-2019.pdf

17, boulevard Raspail
75007 Paris France

+33 (0)1 45 51 09 09
info@ism.edu
www.ism.edu

Analysis on dually listed companies in Hong Kong and China Stock Markets

Thesis

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Olivia Tze Kar CHOI

June 9, 2019

INTERNATIONAL SCHOOL OF MANAGEMENT

Acknowledgements

To my life-coach, my late mother Mrs. Helen Choi: because I owe it all to you. Many Thanks!

With special thanks to Professor Daphne Halkias, Professor Tech Eng and Professor Gianfranco Gianfrate for their fine guidance and unfailing supports for this PhD research thesis. I would also like to thank the research lab staffs at the University of Hong Kong who helped me sorting and dealing with the massive financial data. With a special mention to Professor Cedric Yiu who co-authored my published paper "In-depth analysis in dually listed companies in Hong Kong and Mainland China" in 2013 which also motivated my further studies on the mentioned topic with the Agent-based simulation and limit order book analysis.

And finally, last but by no means least, also to my dear family, Simon and Charlotte, thank you for your caring supports throughout these years.

Table of Contents

Figures & Tables

1 Introduction

1.1 Background and Motivation

Global equity markets are in the midst of technological revolution that redefines the market structure and trading practice due to the rapid growth in electronic communications networks. Financial market microstructure theory, a new research branch in financial market analysis, attracted more attention from researchers (Busato and Handorf, 2001; Nagel 2016).

The new trading platform outperforms the traditional market trading exchanges by providing faster trade execution, lower transaction costs, higher price volatility and tighter spread and deeper market volume (Bakos 1991; Benaroch & Kauffman 2000). The electronic networks provide an alternative platform to conventional equity trading arrangement where trade can be executed within the matching system. When an investor placed an order at a specific price level, electronic network automatically matches and completes the transaction with the counterparty, bypassing the dealer or broker. The automated communication and matching system greatly reduces the trading cost as stocks are exchanged automatically within the network. All outstanding limit buy/sell orders are stored in buy/sell queues that constitute the limit order book. Limit order book is the center of the transaction process in continuous auction market, and it determines the formation of market price of securities (Biais, Hillion, Spatt, 1995).

Due to the complexity of continuous auction market, limit order book modeling has become one of the most important elements in financial market microstructure theory (Chiarella et al 2009). A common practice is to use homogeneous Poisson process to simulate the arrival of new orders, under the assumption of the queuing theory (Cont, Stoikov and Talreja 2010). Agent-based modeling takes a simulation approach to study complex limit order book with dynamically interacting decision makers. Agent-based modeling has been frequently applied in financial markets (LeBaron 2006) and to reproduce stylized facts of the financial system (Palit et al 2012).

Latest research (Chen et al., 2015) on agent based simulation model developed the volatility clustering models to test on the equity price performance and its correlation with future volatilities. Chen et al (2015) observed negative correlation of past price performance with future volatilities in some developed markets, namely the New York stock exchanges.

Over the years, international cross-listing has arisen a great deal of academic focus particularly on the segmented markets theory and the failure of the law of one price in multiple markets. For instance, Gagnon and Karolyi (2010) pointed out the value of international cross-listings: "Cross-listings continue to be vibrant influencing price discovery, trading, and capital-raising for many companies around the world and thus still represent an important force for the integration of global financial markets." A number of studies examine the relationship between price performances and episodes of financial market (Yang et al, 2014; Hui et al, 2017; Chan, 2017) and conclude there is an increasing co-integration of price movement of the cross-listed companies. The existence of long term co-movements of cross-listing company shares performance has importance implication on the Efficient Market Hypothesis as well as portfolio diversification (Taylor & Tonks, 1989; Kasa 1992).

Given the close economic and rapid growth relationships between the Mainland China and Hong Kong, we expect the shares of these dually listed companies to go in tandem. Choi et al [1](2013) found significant price co-integration among these A-shares and H-shares dual-listed stocks between 2006 and 2010. Su, Chong and Yan (2007) found considerable evidences that supported the presence of co-movement of share price cross-listed in Hong Kong and Shanghai between 2002 and 2004.

With this backdrop, this research thesis aims to evaluate the long-term implication and financial value of cross-listing in different stock exchanges based on 1) time series analysis on share prices co-integration, 2) analysis on limit order books

[1] "In-depth analysis in dually listed companies in Hong Kong and Mainland China 2013" focus on evaluating the change in co-integration levels of the stock price prior and posterior to the global financial crisis in 2008. In the paper, an in-depth analysis and discussed from both macro and micro factor analytic perspectives.

employing the stylized LeBaron's agent based model, and 3) investigation on the behavior of market participants and its implications for investment strategy, for major co-listed firms in the Hong Kong Stock Exchange and Shanghai Stock Exchange, respectively. This is a new approach that the analysis is taken from the agent based simulation model for the limit order book and examine the nature of the arrival of orders for cross-listing companies.

In later part of the research, the value of cross listing and the possible theories explaining these implications are to be discussed. The co-integration of price returns is analyzed to investigate the interactions and the behavior of market participants in the stock markets, which is theoretically crucial in understanding of the price dynamics and useful for the optimization of investment portfolio.

1.2 Research Objective and Research Questions

The purpose of this study is to evaluate the cross-listing value and its implications of the co-listed stocks traded on the Hong Kong Stock Exchange and Chinese stock markets, apply a simulated agent based model employing the queuing theory and the Vector Autoregressive Model. So far, research in combination of agent based simulation model and queuing theory on limit order book has been sporadic. Some of the early studies that pursue the idea of explicit analysis for agent based modeling of stock markets using existing order book data by Toke (2015), Panayi et al (2012) and Wang et al (2017). The Vector Autoregressive Model of price returns is analyzed to investigate the interactions and the behavior of market participants, which is useful to understand the structure and formation of the order book and hence derive the price dynamic and inter-correlation between the two exchanges.

A building block quantitative model will be analyzed for the behavior of market participants in these stock markets. As such, it is important to identify suitable questions for the research. In line with the purpose of this study, the central questions are as follows:

1. Given the close economic relationships between Hong Kong and China, shall the same company dually listed in both stock exchanges exhibit same share price performance?

2. Are there any significant differences in asset returns distribution, compositions of market participants and the investment strategies for the dually-listed companies?

3. If the share prices proved to be co-integrated in the long run, what is the implication for portfolio allocation management if investors invested in the same dually listed company in these two financial exchanges?

We attempt to answer these questions by performing some statistical tests on the asset returns, building up the simulated agent-based model and evaluating the levels of stock co-integration from both short-term limit order book and the long-term time series approaches. Hence we conclude the implication for portfolio allocation management if invested in the same stock dually listed in these two financial markets.

1.3 Methodology & Data Description

In this study, major companies dually listed in Hong Kong Stock Exchange and Shanghai Stock Exchange representing the 65% of market capitalization in various sectors including Finance and Banking, Transportation and Logistics, Utilities, Materials, Retails, and Commercial Services. The transaction data including Market Level 1 and Level 2, other financial data including daily share prices, market capitalizations, trading volume, return on assets (ROA), sales growth, and indices data are downloaded from Bloomberg (with additional subscriptions). As A- and H-shares were denominated in RMB and Hong Kong dollars, respectively, all trading data are converted into Hong Kong dollars for analytical purpose.

On a daily basis, every stock recorded more than 4,000,000 transactions from the market opens to closes. Since there is a different market open time between Hong Kong and Shanghai stock exchanges, we only focus on the concurrent trading period for the co-listing companies. The concurrent period is between 9:30am to 11:30am, and 1:00pm-3:00pm.

Day traders or market participants rely heavily on market data, commonly referred as Level 1 market data. The market data includes information about the prices and the completed trades for a market. Traders choose various levels of market depth based on their trading needs. According to worldwide stock exchanges, level 2 market data provides some additional information on the limit order books, including the bid and ask orders currently queuing on the order book. The depth of market, or market depth, is usually referred to the number of contracts (or shares or lots) that are available at each of the bid and ask prices.

In this paper, the market microstructure identifying the relevant aspects of the Hong Kong and Chinese stock markets are analyzed first, followed by study of market dynamics using appropriate agent-based modeling applied on limit order books of the top equities listed on both exchanges.

1.4 Main findings and the contributions to academic literature

Given the close economic relationships between the Mainland China and Hong Kong, we expect the shares of these dually listed companies to go in tandem. Shen et al. (2008) studied the valuation of cross-listings in Hong Kong and China for the period of 1994 – 2003. Su, Chong and Yan (2007) found considerable evidences that supported the presence of co-movement of share price cross-listed in Hong Kong and Shanghai between 2002 and 2004.

This research study expands the valuation on the dually listed companies and tests on the share price co-integration for the period of 2010 - early 2019. We find that 23 out of 35 stocks in our universe show significant share price co-movements and co-integration in the time series of the VECM and Johansen trace test.

We further investigate the microstructural finance analysis and market participants' trading behaviors with the stylized agent-based simulation model (LeBaron 2006).

Using the massive daily order limit book data subscribed from Bloomberg, we conduct an in-depth analysis on trading activities and the share movements.

For the 35 pairs of co-listing companies, we model the limit order book dynamic using Markovian queuing system (Cont et al 2010), and observed that both the arrival rates on limit orders (both bid and ask) and the trading intervals of subsequent orders show positive relationships with the changes in share price, respectively. The mean differences are significant with the t-statistic tests, and the results are noticeable for both Hong Kong and Shanghai markets. It is easily understandable that the trading activities are expected to become more rapid as share price started to move in momentum. We test on the normality of asset returns and find 26 out of 35 stocks report significant leptokurtic fat-tailed distributions, namely the probability of very high or very low returns are higher than that implied by normal distribution.

We calculate the market participant's net inventory and find contradictory conclusion when the change in net holdings is plotted with the changes in share price for Hong Kong. This leads us to examine the capacity of intraday trading activities of traders and share price movements (Hasbrouck and Sofianos 1993). The regression model shows traders' inventory changes are positively correlated to the contemporaneous price changes but negatively correlated with the long-lagged price change in Hong Kong. The regression coefficients for Shanghai data are, however, all positives.

The results reflect that the market participants in Hong Kong and Shanghai, albeit trading on the same company in different exchanges, have very different trading patterns in holding horizons and inventory mean-reversion strategies. Hong Kong financial markets comprised of number of heterogeneous and bounded rational agents, which interact more rapidly through different trading mechanism. The Chinese stock market, on the other hand, is dominated by State-owned-enterprises (SOEs), the government still maintains the position as the largest shareholder, and thus share prices movement is more unified and coherent with the market participant's net holding positions.

In order to find the implications on investment strategy on these dually listed companies based on the different trading patterns, an additional exercise was conducted to compute the aggregate returns of market participants. Our test shows that there exists a significant trading opportunity for China market to catch up with the expected asset returns in their Hong Kong counterparts as the valuation premiums for both exchanges shall emerge when the compositions of market participants, especially in the Shanghai markets, are more diversified and responsive to price movements.

These models are built for further study in price discovery mechanisms, the influence of market microstructure and hence the implication for portfolio allocation management.

1.5 Outline of the Thesis

The rest of this thesis is organized as follows. Chapter 2 reviews the previous literature and theories related to agent-based model and market efficient theory. Chapter 3 describes the historical background and economic development of the Chinese and Hong Kong stock markets. Chapter 4 introduces the scope of research and the quantitative research methodology. Chapter 5 presents the simulation of data using the agent-based models. Chapter 6 is dedicated to the theoretic results where they are explained and discussed. Chapter 7 concludes the study by summarizing the main points as well as suggesting further research directions.

2. Literature Review

2.1.Queuing theory and the limit order book

As early as Stigler (1964) first approach in investigating some rules of the Securities and Exchange Commission on double-auction microstructure, the model attempt to simulate the trading behavior of the market participants. Financial models are built with agents who can exchange shares of stocks according to the defined utility functions of their investment preferences and risk aversion. This belonged to the field of behavioral finance. However, the drawback of these models is its complexity and it is difficult to identify the role of the numerous parameters and types of dependence to these parameters.

Cohen et al. (1985) began to apply queuing theory methods to study the order flow and the limit order book. They assumed the independence of the buy and sell queue and the arrival rate of order remain constant. Domowitz and Wang (1994) proposed a model to characterize the structure of the electronic order book from the perspective of the continuous double auction with price and time priority rules, and analyzed the interaction of the buy queue and sell queue by applying queuing theory with preemptive priorities. With the continuous double auction trading mechanism, they concluded that price volatility and trading volume are higher while the bid-ask spread is usually narrower.

These behavioral models include percolation models (Stauffer, 1998; Cont and Bouchaud, 2000), Ising models (Chowdhury and Stauffer, 1999; Iori, 1999; Kaizoji, 2000; Bornholdt, 2001; Zhou and Sornette, 2007), and minority games (Challet and Zhang, 1997; Challet et al., 2000).

Researchers construct different order-driven models based on different groups of micro driving rules, attempting to simulate the dynamics of the limit-order book (Maslov, 2000; Farmer et al., 2005; Mike and Farmer, 2008; Gu and Zhou, 2009; Tseng et al., 2010). In these models, orders are not specific to strategic investment style, traders are classified and described as zero intelligence traders.

Another important approach for limit order book modeling is that researchers assumed agents are heterogeneous, of which the models are combinations of informed traders, fundamentalists, technical traders, smart traders, and noise traders. Under this framework, one studied the effects of diverse factors, such as noise (Chiarella et al., 2011), technical rules (Chiarella et al., 2009) and investor sentiment (Chiarella et al., 2017).

Cont et al. (2010) transformed the dynamics of limit order book into multi-server queuing system, used a continuous time Markov process to simulate the development of the number of order at each price level, and proposed a continuous time stochastic model of limit order book dynamics. The arrival of limit orders, market orders and cancellation of existing orders are viewed as independent Poisson processes. Cont et al. (2010) postulate that the orders arrive in a unit size measured as the average size of limit orders and that the cancellation rate are proportional to the volume of existing orders. In his research, they calculated some important conditional probabilities that can be used to predict the short-term behavior of limit order book.

Huang and Kercheval (2012) and Cont and de Larrard (2013) put forward two distinctive approaches to estimate the parameters of the queues. Luo (2017) compared the two methods and concluded that the estimators of Cont and de Larrard (2013) model show better fitness using Nasdaq stock market data.

2.2 Agent based simulation models to model investor behavior

Agent based simulation models of financial markets are used to study and understand market dynamics. According to LeBaron (2006), one important aspect in agent based modelling is that prices should emerge internally as a result of trading interactions of market participants. Models in the realm of agent based computational finance view financial markets as inter-acting groups of learning, bounded-rational agents, and agent are heterogeneous and dynamic. Agent based model can enable a complex and adaptive system for research purpose. Chen & Yeh (2002) make a more strict distinction between agent based model and conventional models pointing out that agent based simulation models are composed of many heterogeneous market participants, and prices emerge as a result of interaction among these traders.

A mathematical formalization of agent-based model is suggested in Wooldridge (1999) that the trading environment can be characterized as a set of platform that participants can influence each other through a set of continuous actions. Interaction between agents is a key feature of the agent-based system. Stock prices arise from simulating the interactions of autonomous entities with different profit making strategies. Schelling (1978) used an early agent based work where unexpected aggregate pattern of segregation appears on the top down level. The idea of complexity which emerges from nonlinear interactions between heterogeneous components forms the foundation of Complex Adaptive Systems, which is firstly related to the agent based modelling approach. Computational agent based model can easily accommodate complex learning behavior, asymmetric information, heterogeneous preferences and capture complex interactions between agents (Lovric 2011).

Developing financial models that are capable to mimic the realistic outputs which are qualitatively and quantitatively similar to the empirical data is vital. Lovric (2011) considered one of the first modern agent based models that focus on the link between portfolio insurance strategies and market volatility. He assumed there are two types of investors – re-balancers and portfolio insurers, and the two assets classes in the market, risky stock and cash. Trading time in the model is discrete and the pricing mechanism is based on the order book and portfolio insurance strategy can have a destabilizing effect on the market.

Levy, Levy, Solomon model (LLS) (2000) based on the microscopic simulation approach. The numerical model is developed in the framework of maximization of the expected. The LLS model maps micro level modeling of investor behavior to macro level. With the assumptions of are either Rational Informed Investors (RII) or Efficient market Believers (EMB), the pricing mechanism is determined by the temporary market equilibrium that the total demand for the risky assets equals the total number of outstanding shares.

LeBaron (2006) applied the agent based model that assumes two classes of asset for investments – cash and equity. The agent has a certain wealth and is able to make allocation decision based on their wealth at a discrete time. The model to demonstrate that returns on stocks contain certain characteristics including the distribution of returns is more peaked than the Gaussian distribution, periods of persistent high volatility and correlation between volatility and trading volume.

The queuing modeling methods (Cont et al. 2010) involved strong assumptions in that the order process describing the characteristics of the key parameters have some deviations from the actual market. Recent research (Chen et al. 2015) on agent based simulation model relaxed the assumptions and developed a volatility clustering approach to test on the equity price performance and the correlation with the future volatilities. Liu et al. (1999) and Krawiecki et al. (2002) stated that financial market's volatilities are long-range correlated in time. Chen et al (2015) also showed that the dynamic behavior and community structure of the complex financial market can be characterized by temporal and spatial correlation functions. Jiang et al. (2014) restated that the spatial correlations and sector structure are theoretically crucial in understanding of the price dynamics and useful for the optimization of investment portfolio. Chen et al (2015) found that the correlation between past returns and future volatilities is always negative for the most developed stock markets, namely New York and Hong Kong stock exchanges.

2.3. Theoretical and empirical reasoning behind valuation implication of dual-listing

Efficient market theory emphases if markets were perfectly efficient and asset prices fully reflect all available information, cross listing would be redundant. Investors would be indifference in where to invest and companies would also be indifference in where to list their stocks. In reality, however, there was significant deviation in price performance amongst cross listing stocks in different stock markets and hence arose academic interests in further analysis on valuation implications on cross listing.

Assuming that there were some obstacles in cross-boarder investment, namely the capital control and investment regulation boundary between China markets and Hong Kong, dual-listing could improve the available of stocks for investors that have high barrier to invest. The Capital Market Segmentation Theory suggests that when the investor base increases due to cross listing, risks are shared more widely which eventually lead to lower risk and lower cost of capital (Forster and Karolyi, 1999). In the last section of this research, we aim to conclude the price co-integration as well as the convergence of asset valuations among the cross-listing stocks in China and Hong Kong that would support the Capital Market Segmentation Theory.

3 Historic background and Economic Development in Chinese Equity Markets

Over the last two decades, investment firms and listing companies were motivated to consider whether to offer shares to foreign investors by cross-listing would enhance stock value. The large number of cross listings indicates there are multiply beneficiaries behind the cross-listings. Doidge et al., (2004) stated that cross-listing in foreign stock exchanges show improvement access to larger and deeper pools of capital, and hence increasing market depth, trading volumes and liquidity (Chouinard and D'Souza, 2004).

China and Hong Kong offer an exclusive research framework for cross-listing study, due to the close economic relationship and dependency between the two financial markets. A great number of Chinese companies chose to make their company's initial public offering (IPO) at the Hong Kong Stock Exchange. China stock market is relatively restricted and fully accessible to foreign investors in early years. In this case, it is difficult for international investors to access the rapid growing Chinese equity market and it is almost impossible for Chinese investors to invest in foreign equity market. Hong Kong, on the other hand, tops economic freedom index for 25th consecutive years conducted by the Heritage Foundation and the Wall Street Journal (2012).

3.1 History of Chinese Equity Markets

The Chinese stock markets – Shanghai Stock Exchange ad Shenzhen Stock Exchange were opened in December 1990 and April 11, 1991, respectively. The two markets are relatively young but the pace of growth is very rapid, especially in the last two decades. Thanks to the open economy and the links with Hong Kong, equity markets were developing fast and the market capitalizations were seen to grow from US$3.232 trillion to US$10.687 trillion, a massive three folded growth over the last decade.

Table 3.1 Growth of Chinese Stock Exchanges

	Market Capitalization US$ trillion		No. of listed companies	
	2008	**2018**	**2008**	**2018**
Hong Kong Stock Exchange	1.320	3.936	1,200	2,315
Shanghai Stock Exchange	1.246	4.247	864	1,396
Shenzhen Stock Exchange	0.666	2.504	28	755
Total	**3.232**	**10.687**	**2,092**	**4,466**

It is important to know more about the history of the development of the Chinese equity markets. In the early days of the Chinese equity markets the listed companies could offer only shares that were available for local Chinese investors – A Shares. As the economic started to gain the rapid growing momentum after the Grand Tour of the South by President Deng Xiaoping in late 1980s, the open economy policy intensified international interests in the heated equity market.

The open economy policy initiated in late 1980s also set a path for the listings of state-owned entities (SOEs) and higher importance of the stock market. A couple years later, the first Chinese company, Tsing Tao beer (168 HK), was listed in the Hong Kong Stock Exchange (HKEX) with historic size of market capitalization, which also marked the first H-share listed in Hong Kong and be accessible for International investors.

To further entering the global equity stage, the SOE reforms increased its efficiency and, a few months after the Tsing Tao listing, the Chinese government announced a formal regulation for oversea listings corporations: "A special regulation on raising capital and Listing overseas by a joint stock company.

3.2 Restrictions in China Stock Market

With the strict capital control policy, Chinese investors are prohibited to invest outside of the China equity markets, and foreign investors are also prohibited to invest directly in the Chinese markets in early years. The policy creates multiply problems including liquidity, structural and market segmentation problems. Foreign investors who are qualified with a special status of Qualified Foreign Institutional

Investors (QFII) are allowed to invest in the A-shares. The Chinese government established the QFII policy, which investors can achieve the status of QFII is highly bureaucratic.. The process to applying for the status is required to demonstrate its long-term strategy in China, credibility, capability and commitment second, apply for investment quota. The Chinese authorities are controlling the amount of QFII by hastening the application process. The process of application could last from nine months up to two years. International Institutional Investors have acquired the QFII status since August 2011 has remained the leaders in the markets, namely UBS, Nomura, Goldman Sachs, Government of Singapore Investment Corporations.

Short-selling activities were also banned in the Chinese equity markets until 2010. The China Secretary Regulatory Commission (CSRC) started a trial period for margin trading, short selling and index future in early 2010. Until then, the stock markets in China have grown to a market size of US$ 6.751 trillion (excluding Hong Kong), topping the 3rd place among the international markets. Derivatives trading, short-selling activities are now mature in the Chinese markets, which provide greater liquidity and market transparency for investors for hedging and better management for risks.

Table 1 demonstrates the rapid development in the Chinese stock markets. In terms of IPO, Shenzhen stock exchange has grown rapidly, albeit in a relatively smaller size in market capitalizations. Hong Kong and Shanghai stock exchanges have shown a massive growth in IPOs. In 2010 Hong Kong Stock Exchange raised most capital ranking the first place globally, and followed by Shenzhen Stock Exchange on second place, and Shanghai Stock Exchange came in forth place. Greater China (including Shanghai, Shenshen, Hong Kong and Taiwan stock exchanges) dominated the global number of IPOs (36.5% of global total) and its value of total capital raised is 46.3% globally. (Ernst & Young 2011)

An interesting phenomenon was widely recognized in the academic literature (Lee et al, 2007, Sun et al., 2008 and Peng et al., 2007) that there was a significant difference in the pricing of co-listing stocks in China and Hong Kong, respectively. Seah et al. (2005) highlighted that main issues behind the price gap on A- and H- shares valuation are mainly due to the limited investment opportunity for domestic retail and institutional investors in the mainland China. As supply of A- shares was limited, the Chinese stock markets are highly segmented.

As capital was not allowed to flow freely into and out of the China mainland due to the currency policy, the A- and H- shares were not convertible although each unit represented the same voting right at the listed firm. The A shares lagged the performance of the H- shares in early 2000s as the mainland market was plagued by corporate scandals and insider trading problems. However, as the Chinese government targeted to strengthen its financial markets and to improve market transparency in its economic reform, the stocks market started to soar in mid 2000s.

One distinguished characteristic in China stock market is the higher volatility in its stock trading. The following table shows the volatility in Chinese stock markets as compared to the US in 2018.

Table 3.2: Standard deviation to stock returns (%)

	Shanghai Stock Exchange	Shenzhen Stock Exchange	S&P500
Standard deviation to stock return (%)	17.06%	22.32%	8.17%

In mid-2007, the Hong Kong Hang Seng Index Services Limited launched the Hang Seng China Premium A-H Index to help investors tracking the performance of the major Chinese companies dually listed. The index measured the weighted average premium (or discount) of A-share prices to H-share prices of the index constituents. The higher the index, the higher premium A- share is trading compared to H-shares. The following table shows the dually listed companies in SHSE and SEHK and the shares premium (or discount).

Table 3.3: Dually listed Companies in Hong Kong and Shanghai Stock Exchanges

	Name	Symbol	Price (HKD)	Symbol	Price (RMB)	Premium (%)
1	AIR CHINA	00753.HK	9.11	601111.SH	9.88	-20.83%
2	ANHUIEXPRESSWAY	00995.HK	5.31	600012.SH	6.63	-31.23%
3	BANK OF CHINA	03988.HK	3.72	601988.SH	3.87	-17.47%
4	BANKCOMM	03328.HK	6.65	601328.SH	6.22	-8.20%
5	BEIJING N STAR	00588.HK	2.96	601588.SH	3.89	-34.67%
6	CCB	00939.HK	6.91	601939.SH	7.28	-18.50%
7	CHALCO	02600.HK	3.09	601600.SH	4.19	-36.68%
8	CHINA EAST AIR	00670.HK	5.36	600115.SH	6.65	-30.80%
9	CHINA LIFE	02628.HK	22.05	601628.SH	30.32	-37.56%
10	CHINA RAILWAY	00390.HK	6.08	601390.SH	7.1	-26.48%
11	CHINA SOUTH AIR	01055.HK	6.72	600029.SH	8.17	-29.38%
12	CITIC BANK	00998.HK	5.00	601998.SH	6.28	-31.64%
13	CM BANK	03968.HK	39.55	600036.SH	34.9	-2.70%
14	CONCH CEMENT	00914.HK	47.15	600585.SH	39.39	2.78%
15	COSCO SHIP DEV	02866.HK	1.1	601866.SH	3.35	-71.81%
16	DONGFANG ELEC	01072.HK	5.97	600875.SH	12.29	-58.29%
17	EB SECURITIES	06178.HK	7.73	601788.SH	12.75	-47.95%
18	FIRST TRACTOR	00038.HK	2.03	601038.SH	5.23	-66.67%
19	HUANENG POWER	00902.HK	5.05	600011.SH	6.71	-35.38%
20	ICBC	01398.HK	5.87	601398.SH	5.72	-11.89%
21	JIANGSU EXPRESS	00177.HK	10.78	600377.SH	9.98	-7.26%
22	JIANGXI COPPER	00358.HK	10.62	600362.SH	15.47	-41.06%
23	KUNMING MACHINE	00300.HK	2.49	600806.SH	1.47	45.44%
24	LUOYANG GLASS	01108.HK	2.45	600876.SH	13.26	-84.14%
25	MAANSHAN IRON	00323.HK	3.68	600808.SH	3.56	-11.25%
26	MINSHENG BANK	01988.HK	5.86	600016.SH	6.32	-20.39%
27	NANJING PANDA	00553.HK	3.52	600775.SH	16.17	-81.31%
28	PETROCHINA	00857.HK	5.06	601857.SH	7.45	-41.68%
29	PING AN	02318.HK	92.15	601318.SH	83.95	-5.75%
30	SHENZHENEXPRESS	00548.HK	9.61	600548.SH	10.27	-19.66%
31	SINOPEC CORP	00386.HK	5.97	600028.SH	5.68	-9.76%
32	SINOPEC SSC	01033.HK	1.14	600871.SH	2.85	-65.66%
33	TIANJIN CAPITAL	01065.HK	3.16	600874.SH	9.02	-69.92%
34	TSINGTAO BREW	00168.HK	45.65	600600.SH	48.33	-18.90%
35	YANZHOU COAL	01171.HK	8.27	600188.SH	11.23	-36.77%

Remark: Price as of March 30, 2019

3.3 Characteristics of Efficient Market

Market microstructure studies the mechanism of the price formation in markets, analyzing the processes by which the demand and supply of stocks are translated into transactions and prices. Studying the behavior of the market participants is important to understand the potential impact on the formation of the securities market. According to Fama (1970), a stock market that is considered to be mature should have the following characteristics:

1) Transparency: Market participants expect timely and accuracy of the market information on prices and volumes traded, and on the outstanding bids and asks in order to determine the fair market price with the demands.
2) Liquidity: Liquidity refers to the degree of a stock that can be quickly traded in the market at a fairly certain price.
3) Well organization: The stock market is ensured an orderly market for trading of securities under the regulatory organization. Market participants are not restricted to participate in the market trading activities, and stocks are freely traded. Regulation and policy are clearly stated and not changed unexpectedly.
4) Low transaction costs: Market participants consider transaction cost as the barrier in investing in the stock market. Reasonable transaction cost attracts market participants and increase liquidity.
5) Informational efficiency: Price performance of stock reflects all available information. The major driving force of market dynamic is information.

An efficient market is defined as a market in which prices always fully reflect all available information, according to the Efficient Market Theory. In modern finance theories, the market dynamics were analyzed applying mathematical tools including the portfolio optimization theory by Markowtz, the Capital Asset Pricing Model by Sharpe, the Expected Utility Theory and the Efficient Market Theory. All the models assume that all traders are rational, markets are considered efficient and market prices are determined at a certain equilibrium according to demands and supplies.

With this backdrop, given the close economic and political links between Hong Kong and China, the dually listed companies in the two exchanges are expected to experience consistent price movement. Researchers found significant price co-integration among these A-shares and H-shares dual-listed stocks. A number of studies examine the relationship between price performances and episodes of financial market (Yang et al, 2014; Hui et al, 2017; Chan, 2017) and conclude there is increasing co-integration of price movement of the cross-listed companies. In order to understand and explain the dynamics of the two financial markets, statistical analysis and the agent-based microscopic models are employed in this research paper.

4 Research Methodology

Several approaches exist to study to understand market structure and dynamics. Theoretical studies aim to find explanation through analytical models; empirical research analyses historical data to find correspondence between various factors; experimental studies focus on analyzing the trading behavior and its consequences on the market dynamics. In this paper, the market microstructure identifying the relevant aspects of the Hong Kong and Chinese stock markets are analyzed first, followed by study of market dynamics using appropriate agent-based modeling applied on limit order books of the top equities listed on both exchanges.

4.1 Data Description

In this study, thirty-five companies, constituted for 65% in total market capitalizations, cross-listed in Hong Kong Stock Exchange and Shanghai Stock Exchange representing the significant sectors including Finance and Banking, Transportation and Logistics, Utilities, Materials, Retails, and Commercial Services are examined. The transaction data including Market Level 1 and Level 2 data are obtained from the subscription of the BPIPE Bloomberg's services which provided up to 10 levels of brokers' order. Other financial data including daily share prices, market capitalizations, trading volume and fundamental data are downloaded directly from Bloomberg. To adjust for the currency effects, all trading data are converted into Hong Kong dollars for analytical purpose.

Day traders or market participants rely heavily on market data, commonly referred as Level 1 market data. The market data includes information about the prices and the completed trades for a market. Traders choose various levels of market depth based on their trading needs. According to worldwide stock exchanges, level 2 market data provides some additional information on the limit order books, including the bid and ask orders currently queuing on the order book. The depth of market, or market depth, is usually referred to the number of contracts (or shares or lots) that are available at each of the bid and ask prices.

Level 1 Market Data

- Bid price: The highest price that a trader is willing to buy an asset at.
- Bid size: The number of shares, forex lots or contracts that are available at the bid price.
- Ask price: The lowest price that a trader is willing to sell an asset at.
- Ask size: The number of shares, forex lots or contracts that are available at the ask price.
- Last price: The price at which the most recent trade was completed.
- Last size: The number of shares, forex lots or contracts that were traded in the most recent trade.

Level 2 Market Data

- Highest bid prices: The highest 5 to 15 prices (depending upon the market) where traders are willing to buy an asset. Traders not only see the current bid, but also all the bids currently below it. In actively traded stocks, there will typically be bids every $0.01 below the current bid, and in actively traded futures, there will typically be a bid each tick below the current bid. If there is a gap between the current bid and next bid, that typically means the stock or contract may experience a larger bid/ask spread and less volume.
- Bid sizes: The number of shares, forex lots or contracts that are available at each of the bid prices.
- Lowest ask prices: The lowest 5 to 15 prices (depending upon the market) where traders are willing to sell an asset. This means you not only see the current ask, but also all the asks above the current ask. In actively traded stocks, there will typically be asks every $0.01 above the current ask, and in actively traded futures, there will typically be an ask each tick above the current ask. If there is a gap between the current ask and next ask, that typically means the stock or contract may experience a larger bid/ask spread and less volume.
- Ask sizes: The number of shares, forex lots or contracts that are available at each of the ask prices.

Source: Bloomberg

5 Simulation of Agent-Based Model

5.1 Simulation framework

The purpose of this research study is to apply a simulated agent based model employing the queuing theory on the co-listed stocks traded on the Hong Kong Stock Exchange and the greater China markets. Some of the early studies that pursue the idea of explicit analysis for agent based modeling of stock markets using existing order book data by Toke (2015), Panayi et al (2012) and Wang et al (2017). Prior to conduct the theoretical analysis in the next section, we perform a simulation on the arrival of order queues based on model suggested by Cont et al (2010).

Generally, limit orders submitted may not be traded immediately until they become the best quote and wait for the preceding arrival of an market order, while market order can fulfill the trader's demand instantly. The difference between bid and ask price is called the bid-ask spread. Since one tick is defined as the unit of offer price on the limit order book (1 tick is equal to HKD 0.05), bid-ask spread can be measured by the number of ticks between bid and ask price.

Cont et al (2010) first proposed the idea of modeling the limit order book dynamic using Markovian queuing system. They viewed the submission of a limit order, market order and cancellation order as independent Poisson process. The arrival rate of limit order at each price level depends on the distance from the opposite best quote, while the cancellation rate is proportional to the queue size and a constant market order arrival rate. All types of orders arrive and leave the queue with the same volume that equal to the average size of the limit orders.

Cont et al. (2010) considered a market in which limit orders can be placed on a price grid $\{1, \dots, n\}$ representing each price level. The state of order book can be model with a continuous-time process, denoted by

$$X(t) = \left(X_1(t), \dots, X_n(t)\right)$$

$|X_p(t)|$ is defined as the number of outstanding limit orders at price $p, 1 \leq p \leq n$.

when $X_p(t) < 0, -X_p(t)$ represents the bid quote on the buy side

when $X_p(t) > 0,\ X_p(t)$ represents the ask quote on the sell side

Under this assumption, the ask price which is the lowest quote on the sell side at time t is defined as

$$P_a(t) = \inf\{p = 1, \dots, n, X_p(t) > 0\}$$

Analogously, the bid price which is the highest quote on the buy side at time t is defined as:

$$P_b(t) = \sup\{p = 1, \dots, n, X_p(t) < 0\}$$

Figure 5.1: Sample of Order Book of China Life (2628 HK)

China Life (2628 HK)				10:30:01 HKT
Day High	20.85	Day Low	20.8	June 29, 2017
Net Change	-0.05	Last Trade	20.85	
Total Value	400k	Total Volume	350k	

Bid Price	Quantity (Queue Size) Volume	Ask Price	Quantity (Queue Size) Volume
20.80	6000(5)	20.85	12500(8)
20.75	4000 (5)	20.90	9800(2)
20.70	23400 (7)	20.95	28200(7)
20.65	15300(8)	21.00	26000(5)
20.60	24700(10)	21.05	81000(4)
20.55	21000(3)	21.10	106600(6)
20.50	2100 (4)	21.15	1900(2)
20.45	92000(4)	21.20	2300(4)
20.40	2000(4)	21.25	108300(1)
20.35	6300(5)	21.30	0(0)

Considering the fact that most of the trading events happen in the vicinity of the bid an ask price, for the purpose of keeping track of the number of outstanding orders at a given distance from the bid or ask price level, the number of buy limit orders at distance i from the ask price is defined as:

$$Q_i^B(t) = \begin{cases} X_{P_a(t)-i(t)} & 0 < i < p_a(t) \\ 0 & P_{a\,(t)} \leq i < n \end{cases}$$

The number of sell limits at distance i from the bid price is defined as:

$$Q_i^A(t) = \begin{cases} X_{P_b(t)+i(t)} & 0 < i < n - p_b(t) \\ 0 & n - P_{b\,(t)} \leq i < n \end{cases}$$

Cont et al (2010) used state $x \in Z^n$ and $1 \leq p \leq n$ to represent the evolution of the order book driven by order flows:

$$x_{p\pm 1} = x \pm (0, \dots, 1, \dots 0)$$

The order flows of limit order, market order and cancellation order at each price level are counting process that can be modeled using independent Poisson process.

- Limit buy (sell) orders arrive at a distance i from opposite best quote with a rate of $\lambda^L(i)$
- Market buy (or sell) orders arrive at ask (or bid) price with a rate of $\lambda^M(i)$
- Cancellation of buy (or sell) limit orders at a distance i from the opposite best quote with a rate $\lambda^C(i)\left|Q_i^{B(t)}\right|$, which is proportional to the queue size $\left|Q_i^{B(t)}\right|$. The assumption is that each order can be cancelled with a rate $\lambda^C(i)$, cancellation orders are mutually independent, thus the total cancellation rate for this queue is $\lambda^C(i)\left|Q_i^{B(t)}\right|$.

With these assumptions, the limit order book is viewed as continuous-time Markov chain with state.

Cont et al. (2010) provided a simple estimation method of the rate parameters. If the observed total number of limit orders at a distance I from the best opposite quote in the sample period T is $N^L(i)$, the estimator of $\lambda^L(i)$ is defined as:

$$\lambda^L(i) = \frac{N^L(i)}{T}$$

In Cont et al.'s model, all types of orders arrive or leave the queue with the same average size of the observed limit order. Denote the average size of limit orders by S^L, that of market order by S^M, and that of cancellation by S^C. The arrival rate of market orders is estimated by the ratio of total observed market order N^M relative to sample period T and scaled with S^M/S^L. The rates of market order and cancellation orders are defined as below, respectively

$$\lambda^M(i) = \frac{N^M(i)}{T}\frac{S^M}{S^L}$$
$$\lambda^C(i) = \frac{N^C(i)}{TQ^i}\frac{S^C}{S^L}$$

Cont et al. (2010) provided the methodology for computing the conditional probabilities of various events based on the model framework. This study will focus on the model framework and hence to establish a market simulator for the order execution.

Different from Cont's model using full price grid in the limit order book state, Abergel and Jedidi (2013) proposed the zero-intelligence model using Markovian point process to describe the order flow. The model setup contained a moving reference frame that records the existing volume on both bid/ask side and a boundary condition assumption. They also extended the model by allowing random size of orders and studied the relationship between the shape of book and size of incoming orders. We adopted a finite moving frame in only describing the state of K price levels on each side with notation:

$$(\boldsymbol{a}; \boldsymbol{b}) = (a_1, \dots, a_k; b_1, \dots b_k)$$

where a_i(or b_i) is number of shares on the sell side (or buy) at price level i ticks away from the bid (or ask) price. This representation of limit order book refers to the visible price levels as seen by traders on the screen. The cumulative depth on each side of the ask and bid is defined as,

$$A_i = \sum_{k=1}^{i} a_k \, ; B_i = \sum_{k=1}^{i} b_k$$

The bid price P_b and ask price P_a are the first non-empty price level in the moving frame of buy and sell side, respectively.
According to Abergel and Jedidi's model, we define the estimator of limit order rate and market order rate as:

$$\lambda^M(i) = \frac{N^M(i)}{T};\ \lambda^L(i) = \frac{N^L(i)}{T};\ \lambda^C(i) = \frac{N^C(i)}{TQ(i)}$$

given

$T\ =\ length\ of\ sample\ period$

$N^M(i)\ =\ total\ number\ of\ buy\ market\ trades$

$N_b^L(i) =\ total\ number\ of\ limit\ buy\ order$

$Q(i)\ is\ the\ average\ number\ of\ shares$

submitted at price level i ticks from the opposite best quote

5.2 Simulation Results

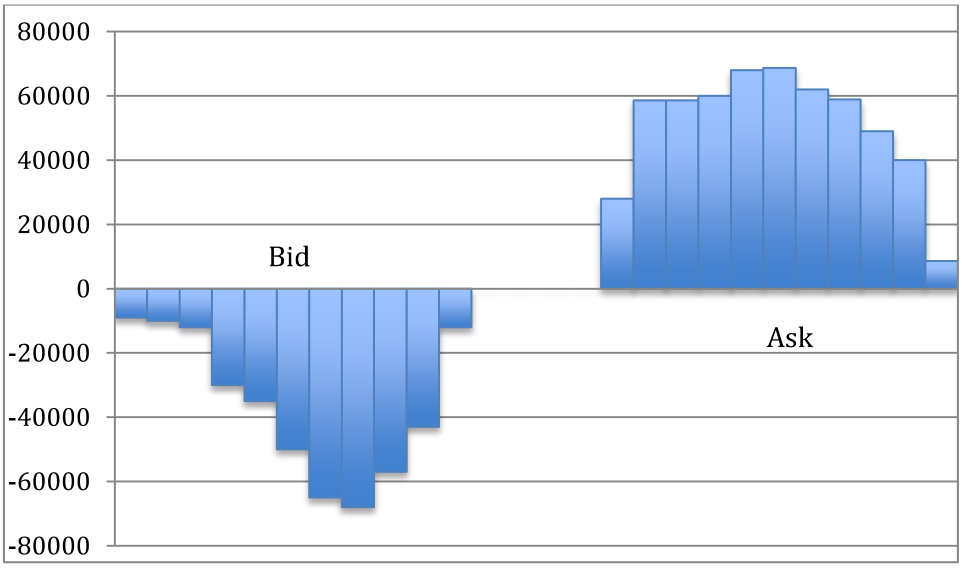
The algorithm can help simulating trading market and provides realistic simulated data to compare with empirical data. The R programming code is used for the simulation as the R language is widely used among statisticians and data miners for developing statistical software and data analysis. (see appendix for code)

High-level data with limit order, market order and cancellation order in quotes and number of shares were used for China Life. (2628 HK) for 10 trading days in May 2019. All levels of transactions were recorded every 0.01 second during the trading hours in Hong Kong; namely there were more than 500,000 records on May 7, 2019. After counting the order frequency and the average market depth of two sides, we observed the following numbers of orders and the limit order book is asymmetric.

Table 5.2 Total number of market, limit and cancellation at each side

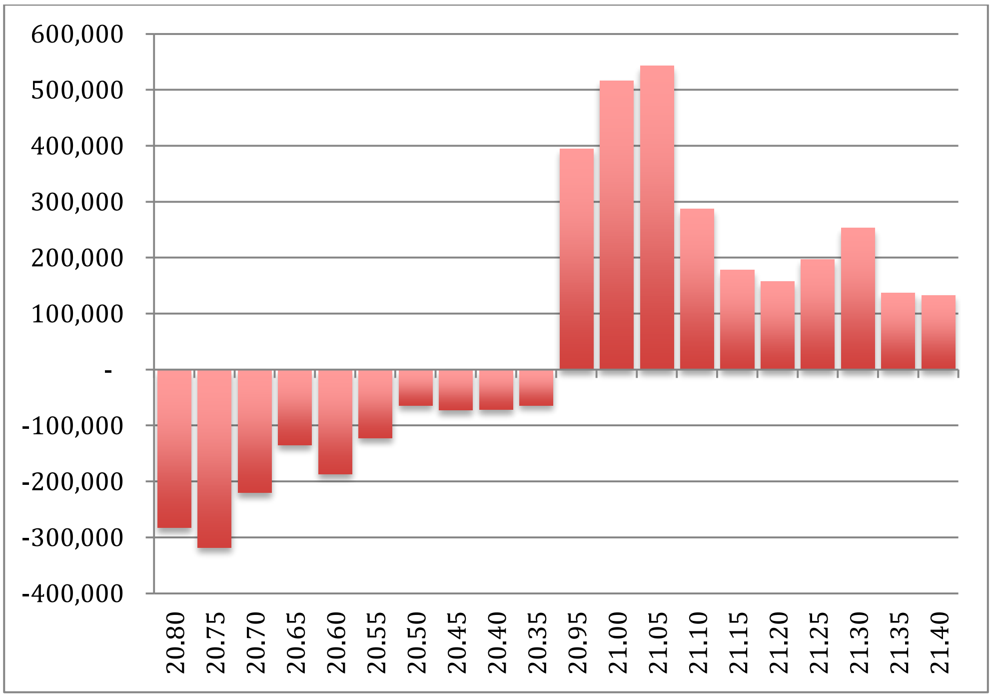
Total Number	Market Order	Limit Order	Cancellation
Buy	31,088	116,714	119,076
Sell	20,953	138,446	100,834

Figure 5.3 Average market depth on both sides



We set the price level limit K of moving frame to be 10 since the high frequency dataset records the available quantity at best 10 quote price levels on each side. We launched a 2 trading days (same length of time as sample data, approximately 500,000 simulated events) simulation and compare these two indicators obtained from simulations to the empirical data.

Figure 5.4 Simulated limit order book on China Life



At this stage the simulated shape of limit order book on both sides is similar to empirical data, especially the average market depth at the first price level. The model shows our assumptions on the Cont et al. and Abergel and Jedidi's models are valid and presented a good model, which is applicable to describe the market liquidity for further analysis.

6 Theoretical Analyses

6.1 Market co-integration

Stock markets are complex and dynamic, market microstructure literature studies the institutional structure behind price formation in markets, analysing the process by which investor's demands are translated into transactions and prices. Studying the market participants' behaviour in agent-based models is important. Market co-integration analysis, as a methodology often used in study of the complex financial systems, can bridge the gap between the micro level of individual investor and the macro level of aggregate market phenomena.

For that purpose, this section elaborates a study on price co-integration with an aim to describe how co-listing stocks are performing while listed in different markets. We use the Vector Error Correction Models (VECM) introduced by Granger (1981), Engle and Granger (1987), as well as the Johansen Trace Test (Johansen, 1998), to test the levels of co-integration among these companies. The results show a stronger and persistent co-integration among the A- H- share performance in the period 2010 to early 2019. The results provide considerable evidence further approve that the public companies, with identical business models and equity valuation, listing in different stock exchanges shall have the same price performance.

To investigate the strength of the common stochastic trend within the co-listed shares, we employ the concept of co-integration relation with Vector Error Correction Models (VECM) (Granger, 1981; Engle & Granger 1987) and perform the Johansen Trace Test (Johansen, 1988) to evaluate the strength of co-integration relation. Note that it was found in Lee and Tse (1996) that the effect of heteroskedasticity does not seriously affect the test results.

6.1.1 Vector Error Correction Models

We model each pair of the co-listed shares price in form of a VECM (*p-1*)

$$\Delta z_t = \Pi z_{t-1} + \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + \mu + e_t$$

Where $z_t = (Z_{1,t}, Z_{2,t})'$ represent the log price of H and A shares, $\Delta z_t = (z_t - z_{t-1})$ is the first difference of the log price, μ is a (2×1) coefficient vector, Π and Γ_i's are (2×2) coefficient matrices, and $e_t = \left(e_{1,t}, e_{2,t}\right)'$ is an unobservable zero-mean independent white noise error.

By substituting $\Pi = -(I - A_1 - \cdots - A_p)$ and $\Gamma_i = -\left(A_{i+1} + \cdots + A_p\right)$for $i = 1,2, \dots, p-1$, the model can be written as a p^{th} order Vector Autoregressive model $VAR\ (p)$ as below,

$$z_t = A_1 z_{t-1} + \cdots + A_p z_{t-p} + \mu + e_t$$

Assuming the process z_t is stationary after taking the first difference then all terms in above equations are $I(0)$ except Πz_{t-1}. In this case, Πz_{t-1} must be $I(0)$ as well. Hence, it contains linear combinations of the variables that are $I(0)$. In this case, Π is singular (i.e.$rank(\Pi) = 1$) containing the co-integration relation and is often referred as the long-run factor, where Γ's are often referred as the short-run factor.

We further imposed the deterministic component, μ, which only affects the long-run factor by restricting $\mu = \Pi \mu_0$, then our VECM can be re-written as

$$\begin{aligned} \Delta z_t &= \Pi(z_{t-1} + \mu_0) + \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + e_t \\ &= \Pi z_{t-1} + \Gamma_1 \Delta z_{t-1} + \cdots + \Gamma_{p-1} \Delta z_{t-p+1} + e_t \end{aligned}$$

Where $\Pi = [\Pi\colon \mu]$ and $z_{t-1} = [z_{t-1}\colon 1]$

The order p is determined by the order specification criteria Akaike's Information Criterion (1974), Hannan-Quinn Criterion (1979), Schwarz Criterion (1978).

6.1.2 Johansen Trace Test

To evaluate the strength of the co-integration relation, we employ the Johansen trace test statistic to test the hypothesis $H_0(r): rank(\Pi) = r$ against $H_1(r): rank(\Pi) > r$ (Johansen, 1994). The Johansen trace test statistic is of the form

$$-T \sum_{j=r+1}^{p} \log(1 - \lambda_j)$$

such that λ_j are the descending ordered eigenvalues solving the generalized eigenvalue problem

$$\det(\lambda S_{11} - S'_{01} S_{00}^{-1} S_{01}) = 0$$

We perform the Johansen trace test with the hypothesis $H_0(r): rank(\Pi) = 0$ and conclude that the H and A shares are cointegrated if the null hypothesis $H_0(0)$ is rejected at the 5% level. Details about the VECM, estimation and testing can be retrieved from Johansen (1995), and Kratzig and Lutkepohl (2004).

If we consider stock co-movement is significant at 1% and 5% levels, we found that 23 out of 35 co-listing stocks under our screen show significant co-integration. It is an exciting finding that majority of the co-listed companies show an emerging shares movement.

Table 6.1: Johansen Trace Test

	Name	H Shares	A Shares	Johansen Trace Test	
1	AIR CHINA	00753.HK	601111.SH	18.022	***
2	ANHUIEXPRESSWAY	00995.HK	600012.SH	19.526	***
3	BANK OF CHINA	03988.HK	601988.SH	18.38	***
4	BANKCOMM	03328.HK	601328.SH	20.805	***
5	BEIJING N STAR	00588.HK	601588.SH	13.489	
6	CCB	00939.HK	601939.SH	13.516	***
7	CHALCO	02600.HK	601600.SH	23.124	***
8	CHINA EAST AIR	00670.HK	600115.SH	36.18	
9	CHINA LIFE	02628.HK	601628.SH	12.136	***
10	CHINA RAILWAY	00390.HK	601390.SH	11.109	***
11	CHINA SOUTH AIR	01055.HK	600029.SH	27.519	**
12	CITIC BANK	00998.HK	601998.SH	20.516	***
13	CM BANK	03968.HK	600036.SH	16.744	
14	CONCH CEMENT	00914.HK	600585.SH	16.502	
15	COSCO SHIP DEV	02866.HK	601866.SH	14.697	**
16	DONGFANG ELEC	01072.HK	600875.SH	25.222	**
17	EB SECURITIES	06178.HK	601788.SH	23.455	
18	FIRST TRACTOR	00038.HK	601038.SH	18.677	
19	HUANENG POWER	00902.HK	600011.SH	25.251	**
20	ICBC	01398.HK	601398.SH	12.079	**
21	JIANGSU EXPRESS	00177.HK	600377.SH	18.681	**
22	JIANGXI COPPER	00358.HK	600362.SH	25.066	**
23	KUNMING MACHINE	00300.HK	600806.SH	14.534	
24	LUOYANG GLASS	01108.HK	600876.SH	23.45	
25	MAANSHAN IRON	00323.HK	600808.SH	13.256	
26	MINSHENG BANK	01988.HK	600016.SH	26.785	**
27	NANJING PANDA	00553.HK	600775.SH	12.345	**
28	PETROCHINA	00857.HK	601857.SH	22.631	**
29	PING AN	02318.HK	601318.SH	19.323	**
30	SHENZHENEXPRESS	00548.HK	600548.SH	21.452	
31	SINOPEC CORP	00386.HK	600028.SH	14.453	**
32	SINOPEC SSC	01033.HK	600871.SH	23.571	***
33	TIANJIN CAPITAL	01065.HK	600874.SH	23.24	
34	TSINGTAO BREW	00168.HK	600600.SH	13.451	***
35	YANZHOU COAL	01171.HK	600188.SH	15.518	

*Remark: The rejection of H0: is marked by ** and *** at 5% and 1% significant level, respectively.*

In our study, the co-integration analysis of A and H shares show the price performance of the major Chinese firms in Shanghai stock exchanges and Hong Kong stock exchanges were co-integrated, 23 out of 35 shares reported a significant and stronger co-integration among A- H- shares performance, according to the Johansen trace test.

6.2 Kurtosis statistics

Our next approach is to examine the limit book model and to test the statistical characteristic that is known as the "stylized fact", namely the fail-tailed distribution. Mandelbrot [17] commenced the study of these characteristics for long time scales, and later Engle and Russell [21] concluded these characteristics can also be observed in intra-day data.

This section evaluates the intraday trading prices on these stylized facts for the major 35 pairs (total 70 companies) of dually listed companies in Hong Kong and Shanghai Stock Exchanges, respectively. A leptokurtic distribution means that when plotting the distribution of returns of two subsequent transactions on the stock, the probability of very high or very low returns are higher than that implied by a Gaussian distribution with the same mean and standard deviation. We measure the degree of fail-tailedness using the statistic kurtosis, the 4th moment of the distribution. A normal distribution has Kurtosis of 3, and any distribution with kurtosis higher than 3 is usually called leptokurtic, or fail-tailed.

On a daily basis, every stock recorded more than 4,000,000 transactions from the market opens to closes. Since there is a different market open time between Hong Kong and Shanghai stock exchanges, we only focus on the concurrent trading period for the co-listing companies. The concurrent period is between 9:30am to 11:30am, and 1:00pm-3:00pm.

Table 6.2: Market trading hours: HKSE, Shanghai and Concurrent

	9:15	9:30	11:30		13:00	15:00	16:00
HKSE		open					close
SSE	open			break		close	
Concurrent							

We find these high frequency daily stock returns are leptokurtic with a significant amount of kurtosis. SPSS gives the kurtosis results in the following table.
26 out of 35 stocks show significant kurtosis, meaning the distributions of price returns are flatter than the standard normal distribution. There is a higher

probability of very high or very low returns for the day-trade and a relatively lower chance of the centered expected mean returns.

Table 6.3: Kurtosis statistics

				Kurtosis			
	Name	**H Shares**	**A Shares**	**H Shares**		**A Shares**	
1	AIR CHINA	00753.HK	601111.SH	1.05	**	1.58	**
2	ANHUIEXPRESSWAY	00995.HK	600012.SH	0.44	**	0.77	**
3	BANK OF CHINA	03988.HK	601988.SH	2.12		3.93	
4	BANKCOMM	03328.HK	601328.SH	1.53	**	3.28	
5	BEIJING N STAR	00588.HK	601588.SH	1.16	**	0.75	**
6	CCB	00939.HK	601939.SH	1.41	**	2.37	
7	CHALCO	02600.HK	601600.SH	3.48		1.65	**
8	CHINA EAST AIR	00670.HK	600115.SH	0.68	**	2.64	
9	CHINA LIFE	02628.HK	601628.SH	0.62	**	2.18	**
10	CHINA RAILWAY	00390.HK	601390.SH	0.39	**	2.89	
11	CHINA SOUTH AIR	01055.HK	600029.SH	1.37		2.90	**
12	CITIC BANK	00998.HK	601998.SH	2.19		8.54	
13	CM BANK	03968.HK	600036.SH	-0.11	**	0.35	**
14	CONCH CEMENT	00914.HK	600585.SH	-0.14	**	-0.18	**
15	COSCO SHIP DEV	02866.HK	601866.SH	1.71	**	2.67	
16	DONGFANG ELEC	01072.HK	600875.SH	2.22		1.43	
17	EB SECURITIES	06178.HK	601788.SH	2.58		2.44	
18	FIRST TRACTOR	00038.HK	601038.SH	1.14	**	1.93	**
19	HUANENG POWER	00902.HK	600011.SH	0.20	**	0.22	**
20	ICBC	01398.HK	601398.SH	1.53		5.02	
21	JIANGSU EXPRESS	00177.HK	600377.SH	-0.73	**	-0.27	**
22	JIANGXI COPPER	00358.HK	600362.SH	1.19	**	3.99	
23	KUNMING MACHINE	00300.HK	600806.SH				
24	LUOYANG GLASS	01108.HK	600876.SH	2.11		2.32	**
25	MAANSHAN IRON	00323.HK	600808.SH	0.63	**	2.68	
26	MINSHENG BANK	01988.HK	600016.SH	0.73	**	8.65	
27	NANJING PANDA	00553.HK	600775.SH	2.93		-0.70	**
28	PETROCHINA	00857.HK	601857.SH	1.50	**	4.06	
29	PING AN	02318.HK	601318.SH	1.59	**	2.41	
30	SHENZHENEXPRESS	00548.HK	600548.SH	0.62	**	0.55	**
31	SINOPEC CORP	00386.HK	600028.SH	1.68	**	1.13	**
32	SINOPEC SSC	01033.HK	600871.SH	2.77		75.52	
33	TIANJIN CAPITAL	01065.HK	600874.SH	0.13	**	1.96	**
34	TSINGTAO BREW	00168.HK	600600.SH	0.84	**	1.10	**
35	YANZHOU COAL	01171.HK	600188.SH	2.68		1.13	

*Remark: The rejection of H0: is marked by ** and *** at 5% and 1% significant level, respectively.*

In the table labelled Tests of normality, the given result of the Kolmogorov-Smirnor statistic assessing the normality of the distribution of scores. A non-significant result (sig. value of more than 5%) indicates normality. For significant value is 0 suggest violation of the assumption of normality.

Table 6.4: Test of Normality

				Kolmogorov-Smirnor			
	Name	**H Shares**	**A Shares**	**H Shares**		**A Shares**	
1	AIR CHINA	00753.HK	601111.SH	0.08	**	0.09	**
2	ANHUIEXPRESSWAY	00995.HK	600012.SH	0.08	**	0.09	**
3	BANK OF CHINA	03988.HK	601988.SH	0.14	*	0.09	**
4	BANKCOMM	03328.HK	601328.SH	0.08	*	0.12	*
5	BEIJING N STAR	00588.HK	601588.SH	0.14	**	0.07	**
6	CCB	00939.HK	601939.SH	0.10	*	0.10	*
7	CHALCO	02600.HK	601600.SH	0.09	**	0.08	**
8	CHINA EAST AIR	00670.HK	600115.SH	0.09		0.13	
9	CHINA LIFE	02628.HK	601628.SH	0.11		0.07	**
10	CHINA RAILWAY	00390.HK	601390.SH	0.09	*	0.09	
11	CHINA SOUTH AIR	01055.HK	600029.SH	0.09	*	0.08	**
12	CITIC BANK	00998.HK	601998.SH	0.07	*	0.13	**
13	CM BANK	03968.HK	600036.SH	0.08	*	0.05	*
14	CONCH CEMENT	00914.HK	600585.SH	0.07	**	0.08	
15	COSCO SHIP DEV	02866.HK	601866.SH	0.11		0.11	**
16	DONGFANG ELEC	01072.HK	600875.SH	0.08	*	0.13	*
17	EB SECURITIES	06178.HK	601788.SH	0.08	*	0.11	
18	FIRST TRACTOR	00038.HK	601038.SH	0.07	**	0.09	**
19	HUANENG POWER	00902.HK	600011.SH	0.08	**	0.09	**
20	ICBC	01398.HK	601398.SH	0.12		0.10	*
21	JIANGSU EXPRESS	00177.HK	600377.SH	0.09		0.07	**
22	JIANGXI COPPER	00358.HK	600362.SH	0.07	*	0.10	
23	KUNMING MACHINE	00300.HK	600806.SH	0.01		0.21	
24	LUOYANG GLASS	01108.HK	600876.SH	0.14	*	0.09	
25	MAANSHAN IRON	00323.HK	600808.SH	0.09	**	0.09	**
26	MINSHENG BANK	01988.HK	600016.SH	0.06		0.13	**
27	NANJING PANDA	00553.HK	600775.SH	0.14		0.10	
28	PETROCHINA	00857.HK	601857.SH	0.09	**	0.09	**
29	PING AN	02318.HK	601318.SH	0.07	**	0.08	**
30	SHENZHENEXPRESS	00548.HK	600548.SH	0.07	*	0.06	
31	SINOPEC CORP	00386.HK	600028.SH	0.12	**	0.05	**
32	SINOPEC SSC	01033.HK	600871.SH	0.13	*	0.32	*
33	TIANJIN CAPITAL	01065.HK	600874.SH	0.09	*	0.10	*
34	TSINGTAO BREW	00168.HK	600600.SH	0.12		0.11	**
35	YANZHOU COAL	01171.HK	600188.SH	0.04		0.05	**

*Remark: The rejection of H0: is marked by ** and *** at 5% and 1% significant level, respectively.*

6.3 Queuing system of the limit order books

In order-driven financial market, traders post market and limit orders in the electronic trading system. A limit buy (sell) order is an order that indicates a buying (selling) signal on a security with certain amount at a fixed price. Investors are free to choose both amount and price at their own discretion. Limit orders accumulate as time goes and they are arranged based on arrival time and price level on the limit order book that reveals both supply and demand situation in the market. The trade price is determined with the matched of bid/ask price when the transaction is executed.

The limit order book is assumed as a continuous-time Markov chain, assuming all limit orders are situated on a finite and discrete price grid $\{1,\ldots,n\}$. At each time, $t \geq 0$, the state of limit order book is modeled by a n-dimensional continuous-time process $X(t) = (X1(t),\ldots,Xn(t))$, where $|Xp(t)|$ is the number of limit orders with price p, $1 \leq p \leq n$. Accordingly, the bid price $p_{B(t)}$ and ask price $p_{A(t)}$are determined by

$$p_{B(t)} = sup\,\{p \in \{1,\ldots,n\}|Xp(t) < 0\} \vee 0$$
$$p_{A(t)} = inf\{p \in \{1,\ldots,n\}|Xp(t) > 0\} \wedge (n + 1)$$

First we calculate the arrival of orders prior to the rise of the share price and observed that the rates of arrival of the *bid orders* significantly *shorten*, and trading activities are more rapid in both Hong Kong and Shanghai markets. The rates of arrival for the *ask orders,* on the other hand, are lengthen, especially in Hong Kong market.

Table 6.5 Average arrival times of bid/ask orders prior to price movements

	Hong Kong			Shanghai		
(Seconds)	Daily average	Prior to rise	Prior to fall	Daily average	Prior to rise	Prior to fall
Arrival of bid	1.52	0.65**	2.91*	2.18	1.02**	2.33
Arrival of ask	2.63	3.22**	1.33**	2.67	2. 97	1.32*
Trade interval	8.99	3.03**	3.54**	1.43	0.83*	1.12*

*Remark: The rejection of H0: is marked by ** and *** at 5% and 1% significant level, respectively.*
Overall average for all 35 dually listed companies in Hong Kong and Shanghai stock exchanges, respectively.

In this section, we directly examine the limit order book up to 5 levels and extract the sequence of transactions to see the trader activity around the price change events. A red flag is defined as the trade price increases and a blue flag is tagged when the price decreases during the concurrent trading hours. At the time trade price change direction, we trace backwards to the previous subsequent order and aggregate the difference of the volumes of all open market bid orders and ask orders at previous price level.

Academic theory links liquidity to the net positions of the risk-bearing capacity of different categories of traders. Huang and Wan (2008) develop an equilibrium framework in which market emerge endogenously when a sudden excess of sell orders overwhelms the insufficient demand. Figure 6.4 and figure 6.5 present the net positions and the price levels of the ICBC (601398 SH) and its limit order book (between 10:04:34 and 10:05:25), respectively.

Figure 6.6: ICBC (601398 SH) share price vs. Net positions

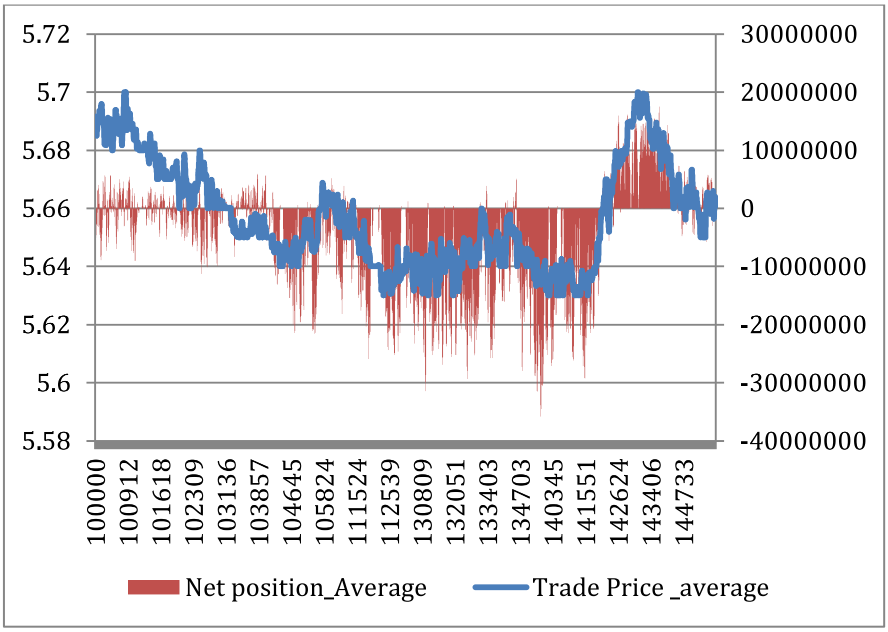
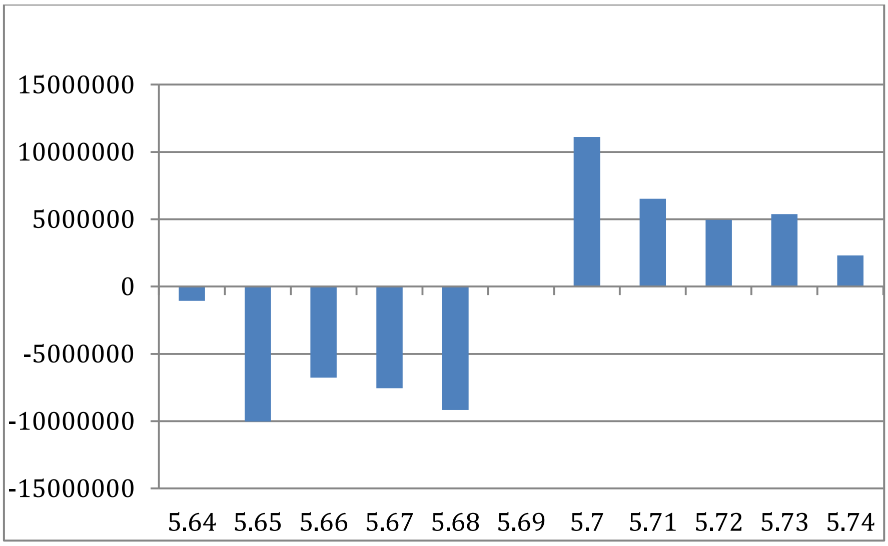
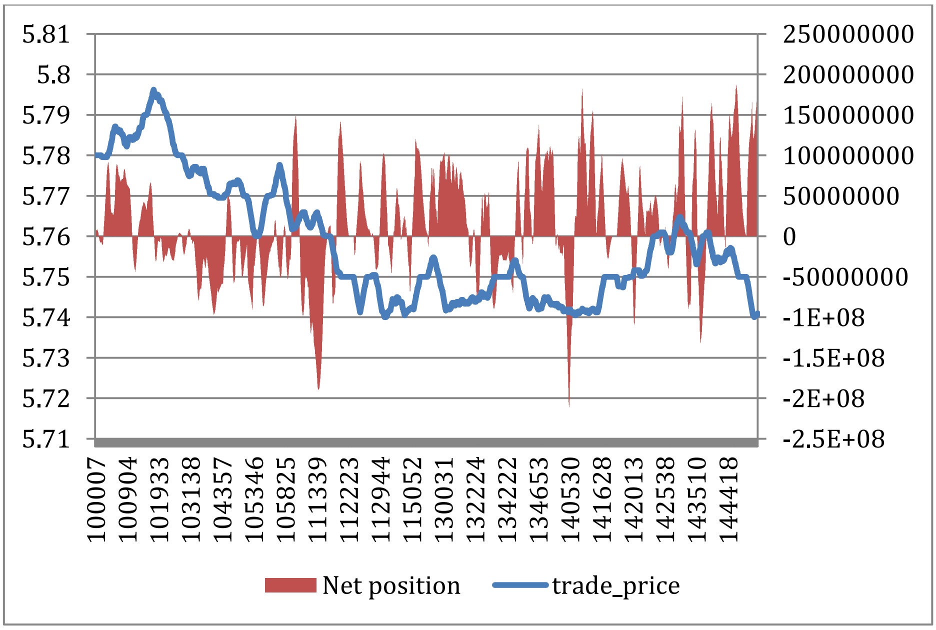


Figure 6.7: ICBC (601398 SH): Limit order book at $5.69



With the limit order book, we observe a major trend in traders' sentiments and share price movements. We find that, for the average 35 stocks listed in Shanghai Stock Exchange, there is a significant relationship between the share price movement and the agent's net positions. The result, however, is not very significant for those 35 listed in Hong Kong. The correlation coefficients for Hong Kong and Shanghai are xx and xx, respectively.

Figure 6.8: ICBC (1398 HK) share price vs. Net positions



6.4 Inventory changes of traders versus price movement

This leads us to examine the capacity of intraday trading activities and the inventory changes of agents in the Hong Kong and Shanghai stock exchanges. Hasbrouck and Sofianos (1993) estimate vector autoregressions that include price changes, signed orders, and the specialist inventory positions for the NYSE. Hendershott and Menkveld (2014) evaluated the price dynamic and inventory performance between specialists in NYSE and NASDAQ.

We attempt to employ a similar test for the statistical relationship between the changes in inventories and changes in prices over the dually listed companies in Shanghai and Hong Kong stock markets, respectively. With this analysis, we examine the co-movement of intraday intermediary inventories and price changes on the following

$$\Delta y_t = \alpha + \phi\, \Delta y_{t-1} + \sum_{i=0}^{20} [\beta_i\, \Delta p_{t-i}] + \epsilon_t$$

Where Δy_t denotes the changes in inventories of trading activities for each second during a trading day, $t = 0$ corresponds to the concurrent trading time on both exchanges from 9:30:00am to t = 28540 at 15:00:00, and Δp_{t-1} denotes the price change between the high-low midpoint of second $t - 1$ and the high-low midpoint of second t to account for bid-ask bounce.

Table 6.9: Regression on agent's net holdings and Prices

	$\Delta y_{t,HK}$	$\Delta y_{t,SH}$
$Intercept$	-1.64	-0.53
Δy_{t-1}	-0.11	-0.03
Δp_{t}	32.09	13.55
Δp_{t-1}	16.28	1.22
Δp_{t-2}	8.36	2.16
Δp_{t-3}	5.09	2.59
Δp_{t-4}	3.91	3.14
Δp_{t-5}	-0.08	1.23
Δp_{t-6}	-0.02	0.15
Δp_{t-7}	-0.21	1.25
Δp_{t-8}	-1.22	2.33
Δp_{t-9}	-4.73	4.58
Δp_{t-10}	-3.46	0.12
Δp_{t-11}	-3.26	1.23
Δp_{t-12}	-4.75	4.24
Δp_{t-13}	-2.74	1.85
Δp_{t-14}	-2.21	3.65
Δp_{t-15}	-2.52	1.12
Δp_{t-16}	-4.36	1.02
Δp_{t-17}	-4.21	0.01
Δp_{t-18}	-5.00	0.57
Δp_{t-19}	-5.26	0.12
Δp_{t-20}	-1.22	4.22

As shown in table, the regression coefficients on the lagged inventory level are both negative, reflecting the mean-reversion of agent's inventory holdings in the markets. Agents in Hong Kong inventory changes are positively related to contemporaneous and lagged prices up to four lags. By the 5th lagged price change, Hong Kong agent's inventory changes become negatively related to price changes. In contrast, Shanghai agent's inventory changes are positively related to contemporaneous price changes and are generally positively related to all lagged price changes.

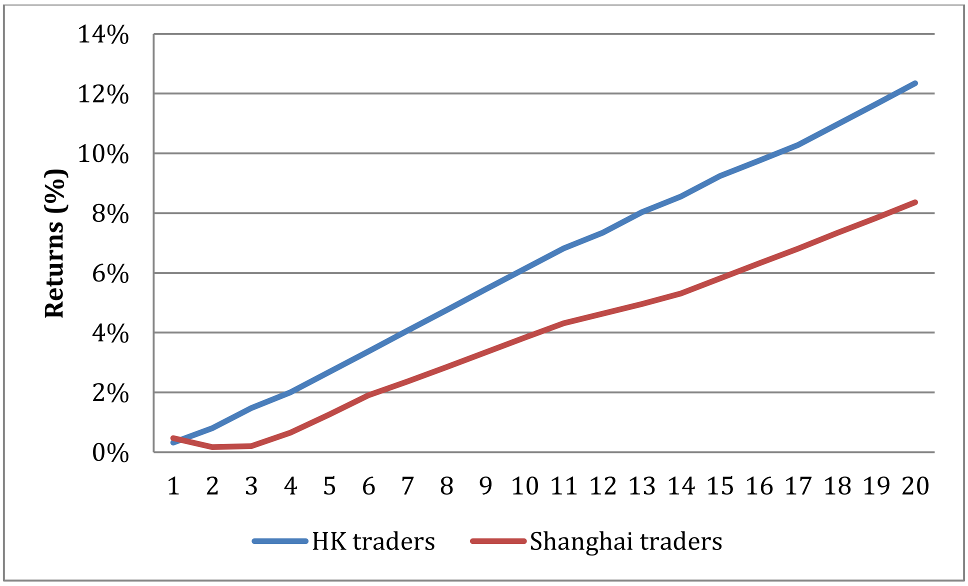
The difference of the regression coefficients of agents in Hong Kong and Shanghai may due to the difference trading patterns in holding horizons and inventory mean-reversion strategies. The market participants are more willing to accommodate trades to longer-time horizon when they are able to trade at a discount relative to future prices in Hong Kong. The Hong Kong agents are more willing to accept buying and selling pressure, which cause price movements that subsequently reverse themselves, implying a positive contemporaneous relationship between inventories and prices.

The result reflects that Hong Kong financial markets mainly comprised of number of heterogeneous and bounded rational agents, which interact through very active trading mechanism. Although we expect the dually listed company's share prices to move in tandem in a long run, based on our previous results from the VECM and Johansen trace test, the analysis on the limit order book and shares movement show a very diverse and different pattern.

6.5 Implications on Trading Strategy

In order to find the implications on investment strategy on these dually listed companies based on the different trading patterns, an additional exercise was conducted to compute the aggregate returns of market participants. We conduct another test to further analyze the statistical relationship between agent's profits and market prices at very short time horizons. Given our first proposition of share price co-integration is approved significant, there exists an arbitrage trading opportunity for co-listing stocks to close the gap in share price premium. We analyze how prices change up to 20 seconds after the agents place orders in Hong Kong and Shanghai markets, respectively. For the "event second" in which agent place an order, average prices paid by the agent in that second are subtracted from the value-weighted average prices for all trades in the same second and each of the following 20 seconds. The results are averaged across event-seconds, weighted by the magnitude of agent's net position in the event-second. Figure 6.10 presents result for the aggregate asset returns for Hong Kong and Shanghai traders for the following 20 seconds period, respectively.

Figure 6.10: Net Asset Returns of traders



In our agent-based models, when Hong Kong traders place orders and trade actively in the stock markets, a higher expected asset return is seen due to the rapid change in inventory holdings in response to the share price movement. This result agrees with the previous regression test of the positive contemporaneous relationships. Shanghai agents report lower aggregate asset returns in the 20-second interval, however.

The results suggest that Hong Kong market participants behave empirically consistent with the theory of limited risk-bearing capacity: they did not take on risky inventories relative to the large and temporary fluctuation of the stock markets, but instead react responsively to the moving momentum of shares. In this case, given the co-integration share performance of dually listed companies, a rapid trading strategy for Chinese stocks shall provide potential arbitrage opportunity to catch up the expected asset returns with their counterparts in Hong Kong. There exists a significant trading opportunity for China market to catch up with the expected asset returns as the valuation premiums for both exchanges shall emerge when the compositions of market participants, especially in the Shanghai markets, are more diversified and responsive to price movements.

7 Conclusion

7.1 Summary

In last two decades, China adopted an open-door policy in economy to strengthen its influence and connection to the world. Hong Kong, being one of the top financial markets worldwide, provided an ideal platform for Chinese firms to raise capital by listing their shares at the Hong Kong Stock Exchange (SEHK).

Given the close economic relationship between China and Hong Kong, this research thesis aims to evaluate the synergy and value-added effects for companies dually listed in SEHK and Shanghai Stock Exchange. Su, Chong and Yan (2007) found considerable evidences that supported the presence of co-movement of share price cross-listed in Hong Kong and Shanghai between 2002 and 2004. The existence of long-term financial co-integration of cross-listed shares has importance implication on Efficient Market Hypothesis (Toda & Philips 1993) as well as portfolio diversification (Taylor & Tonks, 1989; Kasa 1992). International cross-listing has inspired a great deal of academic research over the years particularly on discussion on segmented markets and the failure of the law of one price in those markets.

The new trading platform outperforms the traditional market trading exchanges by providing faster trade execution, lower transaction costs, higher price volatility and tighter spread and deeper market volume (Bakos 1991; Benaroch & Kauffman 2000). Latest research (Chen et al., 2015) on agent based simulation model developed the volatility clustering models to test on the equity price performance and its correlation with future volatilities.

With this backdrop, the purpose of this research thesis is to evaluate the cross-listing value and its implications of the co-listed stocks traded on the Hong Kong Stock Exchange and Chinese stock markets, applying the simulated agent based model with the queuing theory and the Vector Autoregressive Model, respectively. So far, research in combination of agent based simulation model and queuing theory on limit order book has been sporadic.

7.2 Research Results

This research study expands the valuation on the dually listed companies and tests on the share price co-integration for the period of 2010 - early 2019. We find that 23 out of 35 stocks in our universe show significant share price co-movements and co-integration in the time series of the VECM and Johansen trace test. This may be attributed to the rapid economic reform in China and its financial markets have been developing fast, and hence further approved that the public companies, with identical business models and equity valuations, listing in different stock exchanges showed to have the same price performance in a long run.

Market participants in Hong Kong and Shanghai, albeit trading on the same company in different exchanges, have very different trading patterns in holding horizons and inventory mean-reversion strategies. Hong Kong financial markets comprised of number of heterogeneous and bounded rational agents, which interact more rapidly through different trading mechanism. The Chinese stock market, on the other hand, is dominated by State-owned-enterprises (SOEs), the government still maintains the position as the largest shareholder, and thus share prices movement is more unified and coherent with the market participants' net holding positions. As we expect the long-term co-integration of A- H- shares performance, the valuation premium/discount shall emerge if the composition of market participants becomes more diversified and responsive as China continues its rapid economic reform.

We study the intraday trading activities for some major companies listed in both Hong Kong and Shanghai stock exchange, respectively. Our results suggest that the behavior of intraday inventory of traders in Hong Kong and Shanghai react differently in time of temporary buying (selling) pressure. We obtained the second-by-second intraday limit order book from Bloomberg and regress the change in inventory with the change in price and observed that traders in Hong Kong are more willing to accommodate trades to longer-time horizon investors if they are able to buy (sell) at a discount (premium) relative to future prices. Thus the inventories of Hong Kong agents are more willing to accept buying and selling pressure, which cause price movements that subsequently reverse themselves, implying a positive contemporaneous relationship between inventories and prices, and hence a higher expected aggregate asset returns.

In this case, given the co-integration share performance of dually listed companies, a rapid trading strategy for Chinese stocks shall provide potential arbitrage opportunity to catch up the expected asset returns with their counterparts in Hong Kong.

These models are built for further study in price discovery mechanisms, the influence of market microstructure and hence the implication for portfolio allocation management. The study contributes to earlier literature by providing a comprehensive study on the statistical analysis on the share prices performance between Hong Kong and China dually listed companies. In addition, this study covers the current literature in the field of cross-listing. Examining the factors explaining the valuation differences was left outside the scope of this study that poses an interesting area of focus for further research. For some other Chinese companies that are co-listed in US and in Singapore, the same research setting could be further exploited for this ground of study.

Reference

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19(6), 716-723*

Bakos, J. Y. (1991). A strategic analysis of electronic marketplaces. *MIS quarterly*, 295-310.

Benaroch, M. & Kauffman,R. (2000), Justifying Electronic Banking Network Expansion Using Real Options Analysis *MIS Journal*

Biais, B., Hillion, P., Spatt, C., (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *J. Financ. 50, 1655-1689*

Bornholdt, S., (2001) Expectation bubbles in a spin model of markets: Intermittency from frustration across scales. *Int. J. Mod. Phys. C 12, 667-674*

Brock W.A., and Hommes C.H. (1989), Heterogeneous beliefs and routes to chaos in a simple asset pricing model, *Journal of Economic Dynamics and Control, 22(8–9):1235 – 1274.*

Busato & Handorf (2001), The Electronic Communication Networks and the Internet: A New Concept of Business in the Capital Market

Chowdhury, D., Stauffer, D., (1999). A generalized spin model of financial markets. *Eur Phys. J.B. 8, 477-482*

Challet, D., Marsili, M., Zhang, Y.C. (2000). Modeling market mechanism with minority game. *Physica A 276, 284-315*

Challet, D., Zhang, Y.C., (1997) Emergence of cooperation and organization in an evolutionary game. *Physica A 246, 407-418*

Chiarella, C. Iori, G., Perello, J., (2009). The impact of heterogeneous trading rules on the limit order book and order flows. *J. Econ. Dyn. Control 33 (3), 525-537.*

Chiarella, C., He, X.Z., Zhang, M., (2011). An analysis of the effect of noise in a heterogeneous agent financial market model. *J. Econ. Dyn. Control 35 (1), 148C162*

Chiarella, C., He, X.Z., Shi, L., Wei, L. J., (2017). A behavioral model of investor sentiment in limit order markets. *Quant. Financ. 17, 71-86.*

Cont, R., Bouchaud, J.P., (2000). Herd behavior and aggregate fluctuations in financial markets. *Macroecon Dyn. 4, 170-196*

Chen, J.J., Tan, L. & Zheng, B. (2015) Agent-based model with multi-level herding for complex financial systems *Scientific Reports*

Chen S.H., Yeh C.H., Liao C.C. (2002), On AIE-ASM: a software to simulate artificial stock markets with genetic programming, Evolutionary Computation in Economics and Finance, *Physica-Verlag, Heidelberg, pp. 107–122*.

Choi, O.T.K., Wong, H., Yiu, C., Yu, M., (2013), In depth analysis of dually listed companies in Hong Kong and China Stock Markets prior and posterior to the global financial turmoil. *Int'l Journal of Economics and Finance, vol 5. No 10*

Chouinard, Eric. And D'Souza, Chris. (2004). The Rationale for Cross-Border Listings. *Bank of Canada Review 2003-2004*

Cohen, K.J., Conroy, R.M. & Maier S.F. (1985), Order flow and the quality of the market, *Lexington-Heath Books Co.*

Cont, R., Stoikov, S., & Talreja, R., (2010). A Stochastic Model for Order Book Dynamic. *Operations Research, Vol. 58, No.3*

Doidge, Craig., Karolyi, G.A. and Stulz, Rene M. (2004). Why are Foreign Firms Listed in the US Worth More. *Journal of Financial Economics 71, 205-238*

Domowitz I., Wang J.(1994), Auction as algorithms: computerized trade execution and price discovery *Journal of Economic Dynamics and Control*

Gu, G.F., Zhou, W.X., (2009) On the probability distribution of stock returns in the Mike-Farmer model. *Eur. Phys. J.B 67,585-592*

Hasbrouck, J., Sofianos G. (1993). The trades of market makers: An empirical analysis of NYSE specialists, *Journal of Finance 48, 1565-1593*

Hendershott, T. and Menkveld A. (2014). Price pressures, Journal of Financial Economics 114, 405-423

Huang, H., Kercheval, A.N (2012). A generalized birth-death stochastic model for high frequency order book dynamics. *Quant Finance, 12 (4), 547-557.*

Huang, Jennifer, and Jiang Wang (2008), Liquidity and Market Crashes, *Review of Financial Studies 22, 2607-2643*

Jiang, X.F., Chen, T.T. Zheng, B. (2014) Structure of local interactions in complex financial dynamics. *Sci. Rep. 4, 5321*

Kaizoji, T., (2000) Speculative bubbles and crashes in stock markets: An interacting agent model of speculative activity. *Physica A 27, 493-506*

Kirman K. (1993), Ants, rationality, and recruitment, *The Quarterly Journal of Economics,* 108(1):137–156.

Lee, Bong-Soo., Rui, Oliver. And Wu, Wenfeng (2007). Market Segmentation and Stock Prices Discount in the Chinese Stock Market: Revisiting B-share discounts in the Chinese stock market.

Levy M., Levy H, and Solomon S. (2000), Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena, *Academic Press.*

Luckock, H. (2003) A steady-state model of the continuous double auction, *Quantitative Finance, Vol. 3*

Luo, Y.Y. (2017). Statistical Analysis of Markovian Queuing Models of Limit Order books. *Arts & Sciences Electronic Theses and Dissertations. 1076*

Lux T. and Marchesi M. (1999) Scaling and criticality in a stochastic multiagent model of a financial market, *Nature, 397:498–500.*

Liu, Y. (1999), Statistical properties of the volatility of price fluctuations. *Phys Rev. E 60, 1390*

LeBaron, B., (2006) Agent-based computational finance, *Handbook of Computational Economics. vol 2 p. 1187-1233*

Lovric M. (2011), "Behavioral Finance and Agent-Based Artificial Markets", PhD Thesis, RSM Erasmus University/ Erasmus School of Economics.

Iori, G., (1999) Avalanche dynamics and trading friction effects on stock market returns. *Int. j. Mod. Phys. C 10, 1149-1162*

Maslov, S., (2000). Simple model of a limit order-driven market. Physica A 278, 571-578

Mike, S., Farmer, J.D., (2008). An empirical behavioral model of liquidity and volatility. *J. Econ. Dyn. Control 32(1), 200-234*

Nagel, J. (2016) Electronic trading in fixed income markets. *Report submitted by a Study Group established by the Markets Committee. Bank of International Settlement. Publication permalink: www.bis.org.*

Peng, Wensheng, Miao, Hui. And Chow, Nathan. (2007) Price convergence between dual-listed A and H shares. *Hong Kong Monetary Authority – China Economic Issues 6*

Palit, I., Phelps, S. Ng, W.L. (2012) Can a zero-intelligence plus model explain the stylized facts of financial times series data?. *11th International Conference on Autonomous Agents and Multi-agent Systems, p. 653-660*

Panayi E., Harman M., Wetherilt A. (2013) Agent-Based Modelling of Stock Markets Using Existing Order Book Data. *Giardini F., Amblard F. (eds) Multi-Agent-Based Simulation XIII. MABS 2012. vol 7838. Springer, Berlin, Heidelberg*

Reese, William A. J. and Weisbach, Michael S. (2002). Protection of minority shareholder interests, cross-listings in US, and subsequent equity offerings. *Journal of Financial Economics 66, 65-104*

Stauffer, D., (1998), can percolation theory be applied to the stock market? *Ann Phys. 7, 529-538.*

Sun, Qian., Wu, Yujun. And Tong, Wilson H.S. (2008). Why would Chinese Firms List Overseas.

Tseng, J.J., Lin, C.H., Lin, C. T., Wang, S.C., Li, S.P., (2010) Statistical properties of agent-based models in markets with continuous double auction mechanism. *Physica A 389, 1699-1707*

Toda, H. Y., & Phillips, P. C. B. (1993). Vector autoregressions and causality. *Econometrica, 61(6), 1367–1393*

Toke, I.M., (2015), The order book as a queuing system: average depth and influence of the size of limit orders. *Quantitative Finance 2015, 15(5) p.795-808.*

Wang X.T., Wellman, M.P. (2017), Spoofing the limit order book: an agent-based model, *AAMAS '17 Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems P. 651-659*

Wooldridge, M. (1999), Intelligent agents, in G. Weiß, ed., Multiagent Systems: a Modern Approach to Distributed Artificial Intelligence, *MIT Press, Cambridge, Massachusetts, pp. 27–77.*

Zhou, W.X., Sornette, D., (2007). Self-organizing Ising model of financial markets. *Eur Phys. J.B 55, 175-181.*

Appendix A: Definition of Key Terms

Aggregation Gaussianity: The shape of the distribution is not the same for all time scales, as the time scale is increased (returns are calculated over a longer period) the distribution becomes more normal.

Absence of autocorrelations: Autocorrelations in asset returns are often insignificant, except for very high frequency data where microstructure starts playing a role.

Asymmetry in time scales: Coarse-scale volatility measures predict fine scale volatility measures better than the other way. Cont et al (2006) points out that those stylized facts should all be considered as mostly "qualitative properties" of asset returns since they may not be precise enough to distinguish between different quantitative models (for example, many distributions can be used to fit heavy-tailed data)

Autoregressive conditional heteroskedasticity: the autoregressive conditional heteroskedasticity (ARCH) model is a statistical model for time series data that describes the variance of the current error term as a function of the actual sizes of the previous time periods' error terms. ARCH models are commonly employed in modeling financial time series that exhibit time-varying volatility clustering.

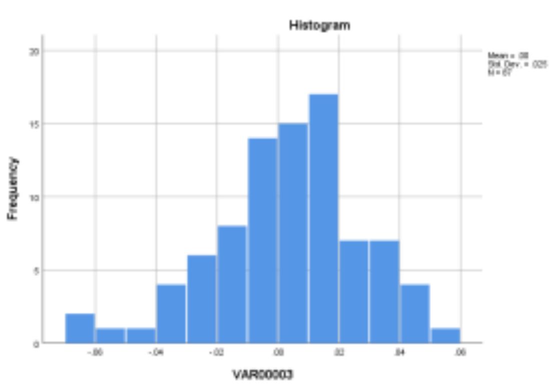
Block order: A limit order with a minimum size of 50 contracts.

Cancellation order: An order to cancel an existing limit order.

Conditional heavy tails: Returns that have been corrected for volatility clustering still exhibit some degree of heavy tails.

Gain/loss asymmetry: It is possible to observe large downward movements in stock prices and stock index values, but not equally large upward movements.

Heavy tails (fat tails): The unconditional distribution of returns tends to be non-Gaussian, sharp peaked and heavy tailed.



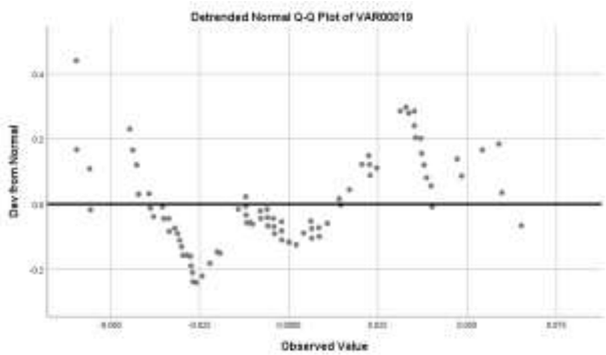
Intermittency: At any time scale, returns display a high degree of variability.

Leverage effect: A negative correlation between the volatility of returns and returns themselves.

Limit order: A limit order is an order to buy or sell a contract at a specified price or better.

Market order: A market order is an order to buy or sell an asset at the bid or offer price currently available in the market place.

Volatility Clustering: refers to the observation, first noted as Mandelbrot (1963), that "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes. Volatility is not constant.



Appendix B Numerical Simulation

The framework and simulation algorithm of model is listed as follow:

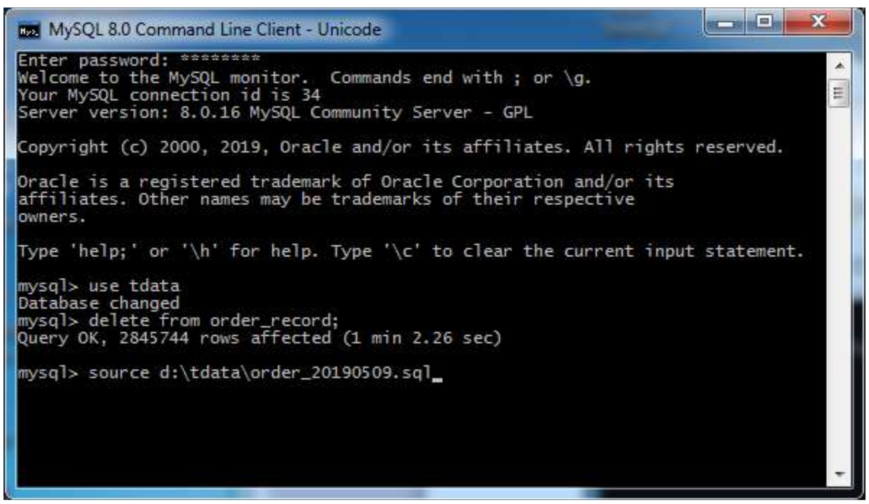
1. Order book on each side is represented by a moving frame with K price levels. Boundary condition is included.
   a. Model parameters defined:
      - Visible price level limit $\boldsymbol{K}$;
      - Order size density function parameters $(\boldsymbol{K^M},\boldsymbol{\alpha^M}),(\boldsymbol{K^L},\boldsymbol{\alpha^L}),(\boldsymbol{K^C},\boldsymbol{\alpha^C})$, initial state of moving frame $(\boldsymbol{a};\boldsymbol{b})$;
      - Initial bid price $\boldsymbol{P_b}$ and ask price $\boldsymbol{P_a}$;
      - Length of interest simulation time $\boldsymbol{T}$;
   b. Initialize volume of full price grid; simulate an event time τ that is exponentially distributed with parameter λ
2. Arrivals of limit order, market order and cancellation are mutually independent Poisson process.
   a. Simulate an event type (e.g. Market order, Limit order, Cancellation order) which follows a discrete distribution with a probability vector;
   b. If event is "Market order arrival", simulate order size using power-law probability density function with parameters $(\boldsymbol{K^M},\boldsymbol{\alpha^M})$, update ***a*** (buy market order case), or ***b*** (sell market order case). Note that market order decreases the availability quantity on non-empty queue from level 1 to level ***K*** until the demand is fulfilled.
   c. If event is "Limit order arrival", simulate order size using power law probability density function with parameters $(\boldsymbol{K^{\square}},\boldsymbol{\alpha^L})$, and determine which price level it arrives at by the probability vector $(\boldsymbol{\lambda_b^L}(\boldsymbol{1}),\dots\boldsymbol{\lambda_b^L}(\boldsymbol{K}))$. Update ***a*** (buy limit order) or ***b*** (sell limit order).
   d. If event is "Cancellation", simulate order size using power law probability density function with parameters $(\boldsymbol{K^C},\boldsymbol{\alpha^C})$, and determine which price level it arrives at by the probability vector $(\boldsymbol{\lambda_b^C}(\boldsymbol{1}),\dots\boldsymbol{\lambda_b^C}(\boldsymbol{K}))$. Update ***a*** (buy case) or ***b*** (sell case).
3. Update parameters and repeat simulation until procedure exceeds length of interest simulation time ***T***.
   a. Update P_b as the price of first non-empty level of b, update P_a as the price of first non-empty level of ***a***.
   b. Update the full price grid similar to step 1 with updated $\{\boldsymbol{P_a},\boldsymbol{P_b},\boldsymbol{a},\boldsymbol{b}\}$ go to step 2.

Appendix C Big Data Analysis, MySQL

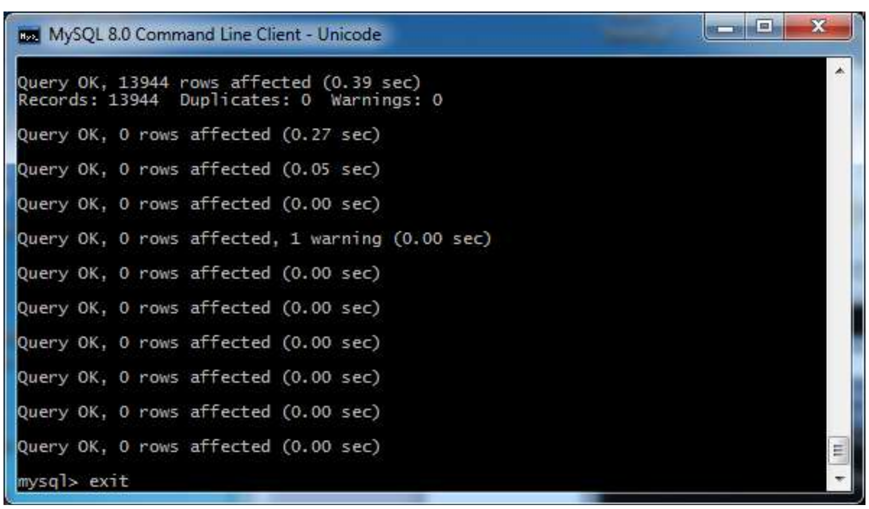
MySQL Workbench will be replaced PHPMyAdmin and I have use it to create the scheme ‘tdata’ and use SQL query to create to the table ‘order_record’.

Every time step:

a. Close any MS Access, MS Excel and MySQL Workbench before import the new data
b. Use ‘7-zip’ to **exact the ‘.sql’ files** into **folder ‘d:\tdata\’** from your newly received ‘.rar’ file.
c. Use another tool: ‘MySQL 8.0 Command Line Client – Unicode’ from desktop to import the ‘SQL’ file directly
d. Open the console
e. Type password: vertrigo (small-letter) to login
f. Type command: **use tdata** (the scheme name)
g. Type command: **delete from order_record;** (clear records from table order_record)
h. Type command: **source d:\tdata\xx.....xx.sql**
i. Type ‘exit’ to exit after import finished



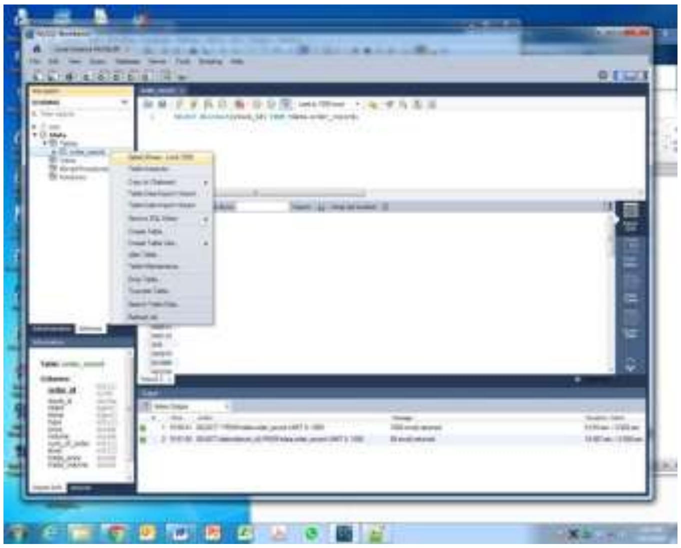

```
Enter password: ********
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 34
Server version: 8.0.16 MySQL Community Server - GPL

Copyright (c) 2000, 2019, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> use tdata
Database changed
mysql> delete from order_record;
Query OK, 2845744 rows affected (1 min 2.26 sec)

mysql> source d:\tdata\order_20190509.sql
```




```
Query OK, 13944 rows affected (0.39 sec)
Records: 13944  Duplicates: 0  Warnings: 0

Query OK, 0 rows affected (0.27 sec)

Query OK, 0 rows affected (0.05 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected, 1 warning (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

Query OK, 0 rows affected (0.00 sec)

mysql> exit
```

Usage of MySQL Workbench:

Open the MySQL Workbench
Select tdata -> order_record and you can type your SQL statement to query
SELECT distinct(stock_id) FROM tdata.order_record;
right-click the 'order_record' and choose 'select row – limit 1000' to list out first 1000 row of the table as shown as below dump screen.



Create a SQL view of each stock ID,

For example the SQL statement of view named stock168 for stock ID=168:

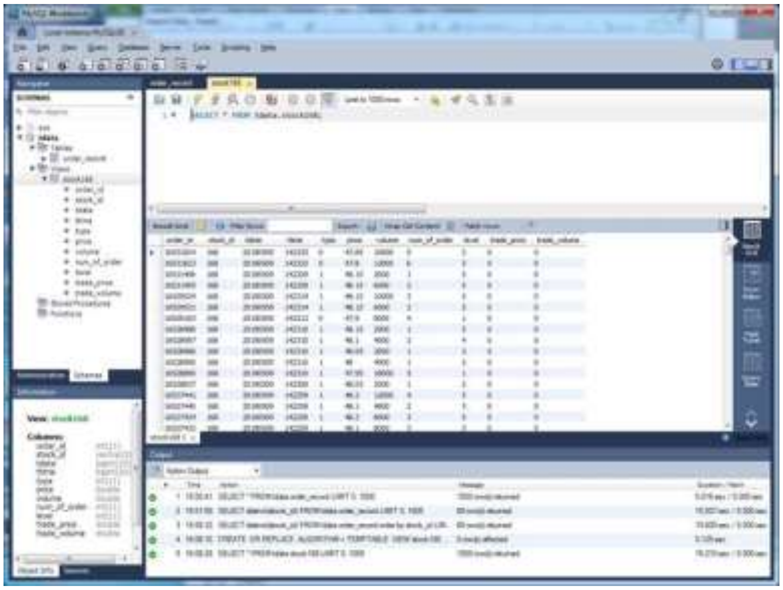
CREATE

ALGORITHM = TEMPTABLE

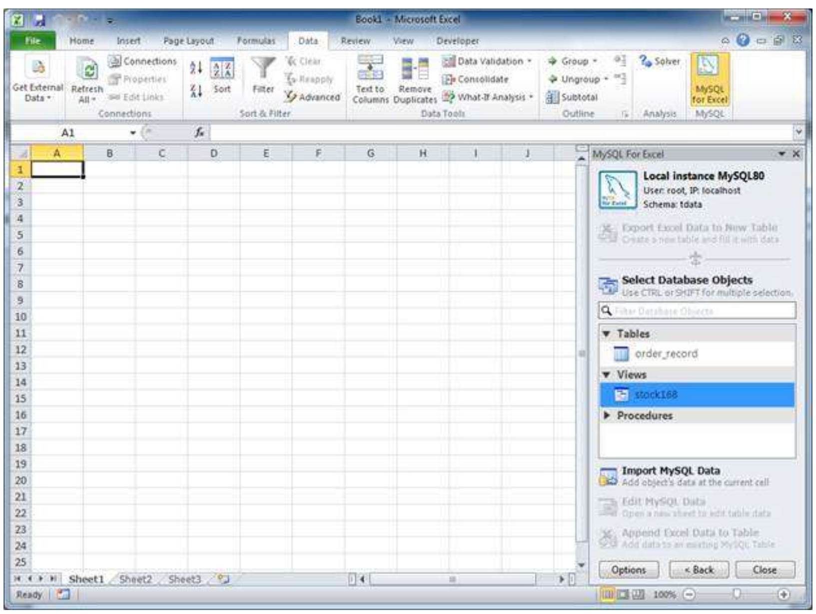
VIEW stock168

AS SELECT * FROM tdata.order_record where stock_id=168;

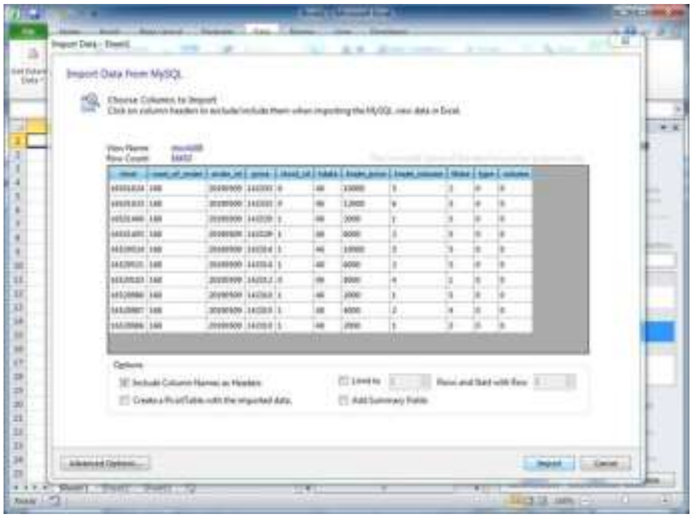
Browse the filtered records of stock id =168 as follows:

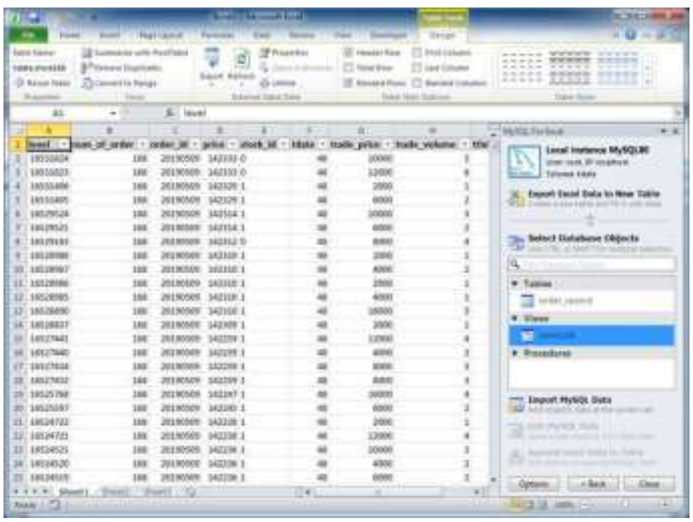


Use ‘MySQL for Excel’ to import the data of that particular stock ID to Excel.



- Select stock168 from Views
- Click Import MySQL Data
- Click import

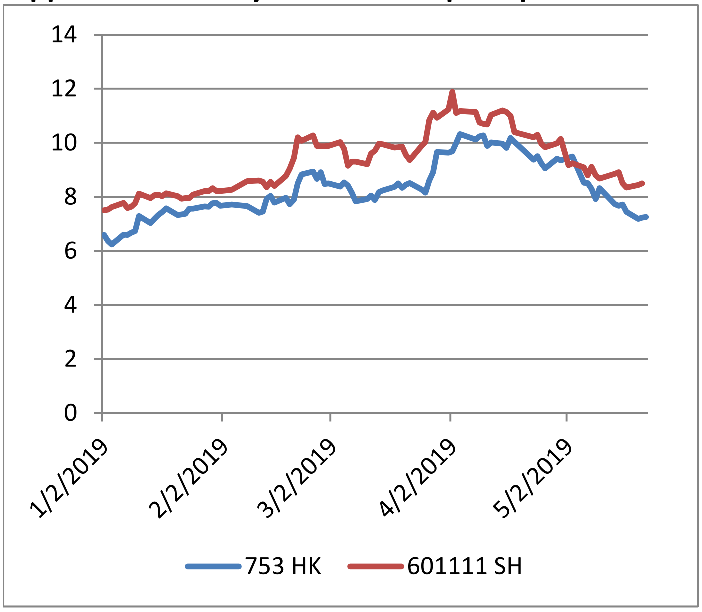


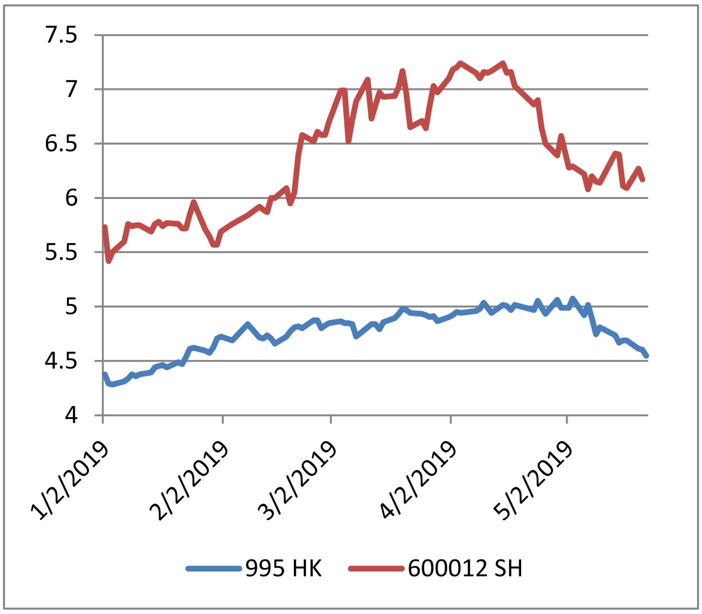


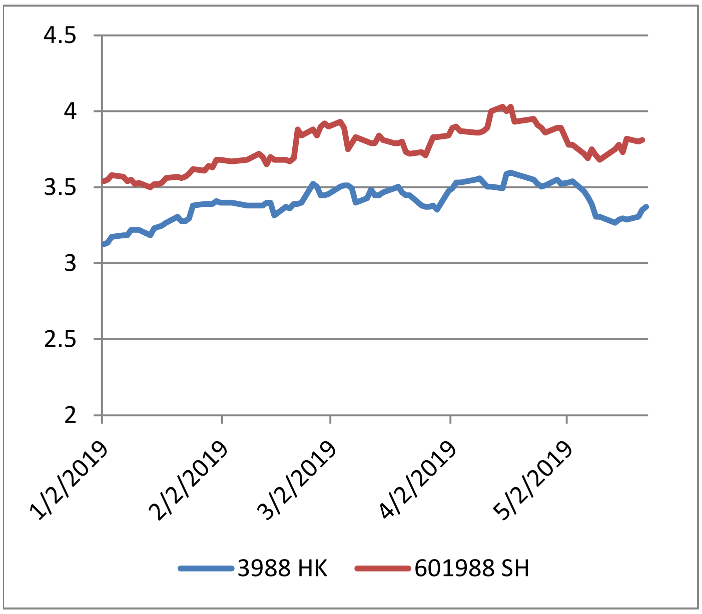
Create different SQL view for each stock ID yourself and you can use those views to retrieve the transaction record of each stock ID in Excel.

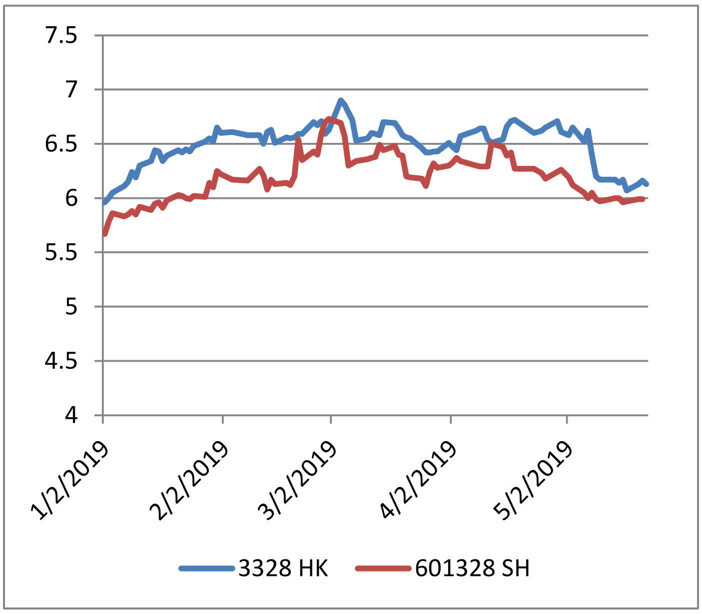
SQL code: CREATE ALGORITHM = TEMPTABLE VIEW stock168_601168 AS SELECT * FROM tdata.order_record where (stock_id=168 or stock_id=601168) and (ttime between 100000 and 150000) order by stock_id, ttime;

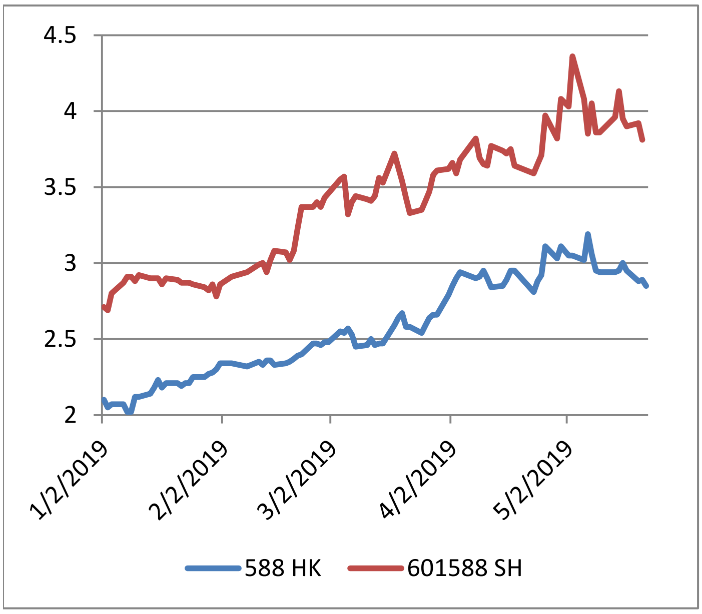
Appendix D: Dually-listed stocks price performance charts

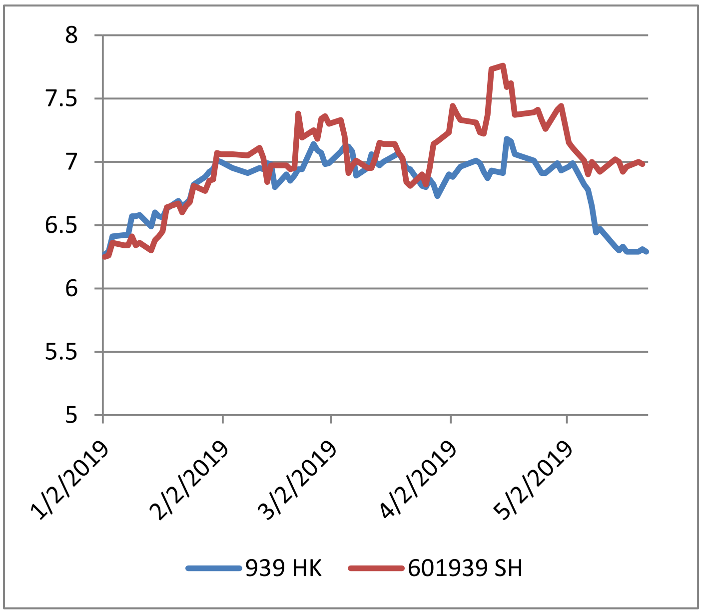


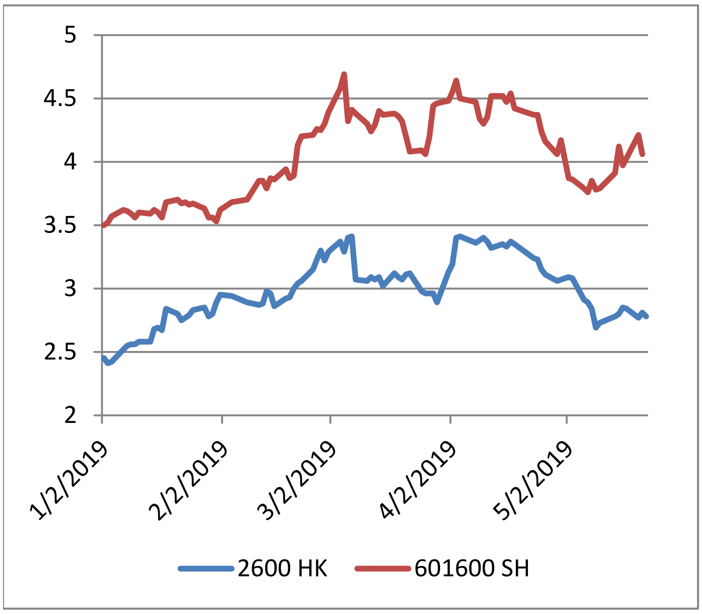




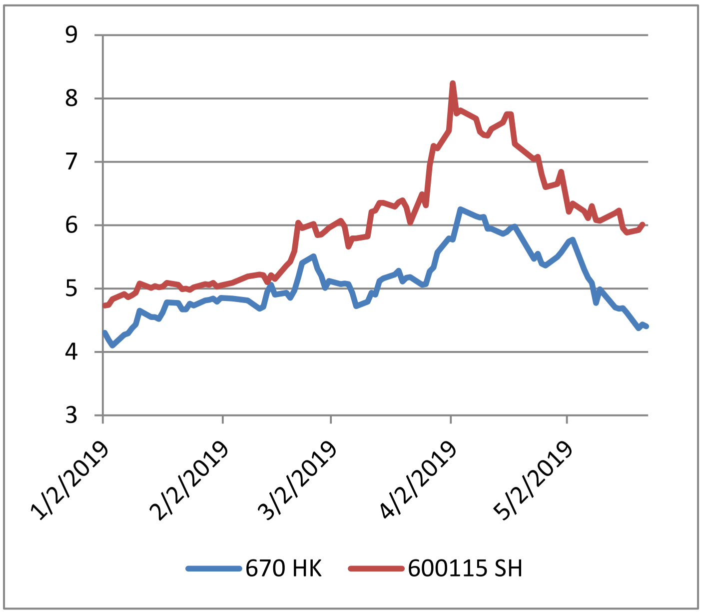




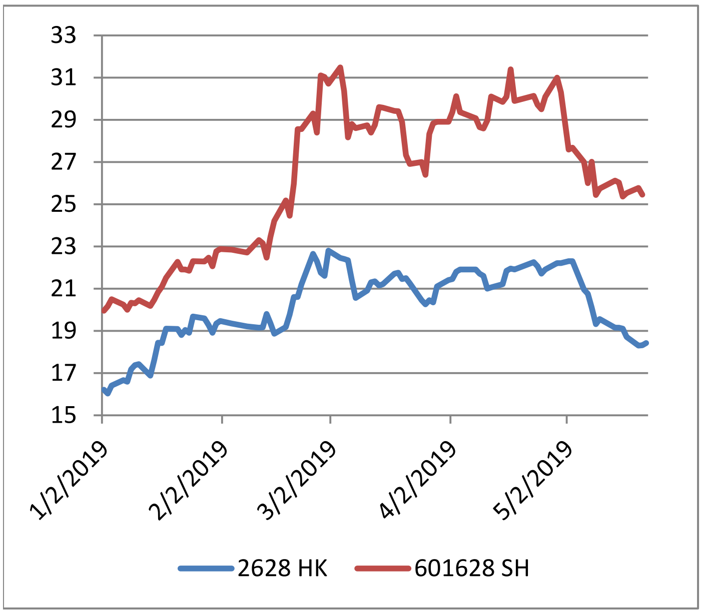




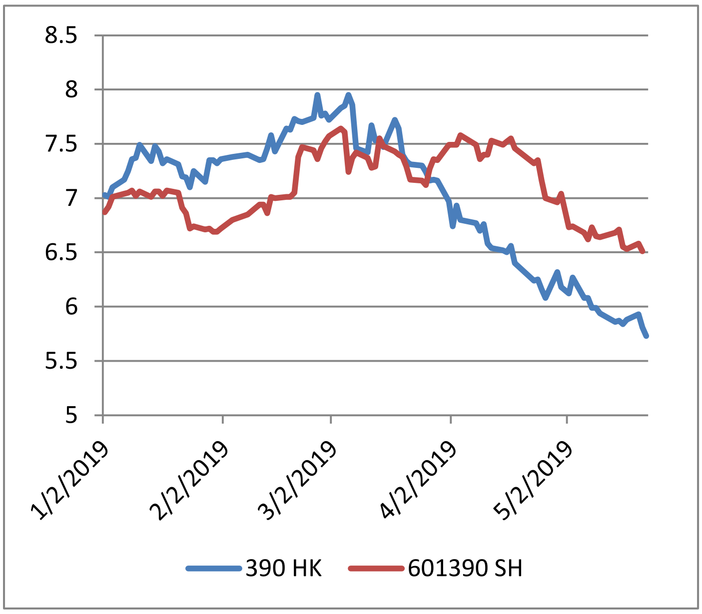
5
4.5
4
3.5
3
2.5
2
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
2600 HK
601600 SH



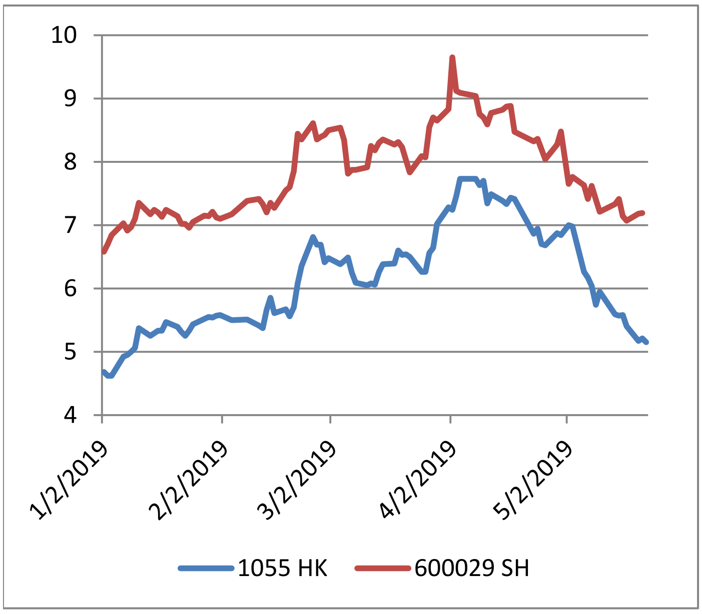
9
8
7
6
5
4
3
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
670 HK
600115 SH



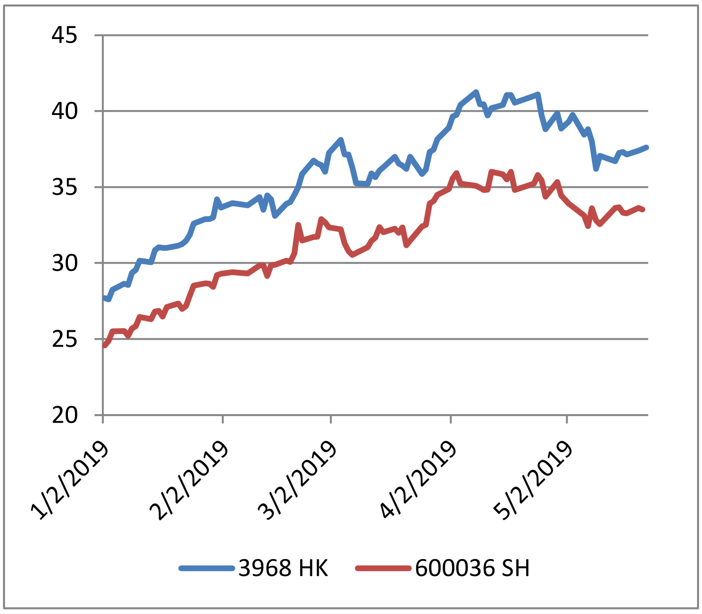
33
31
29
27
25
23
21
19
17
15
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
2628 HK
601628 SH



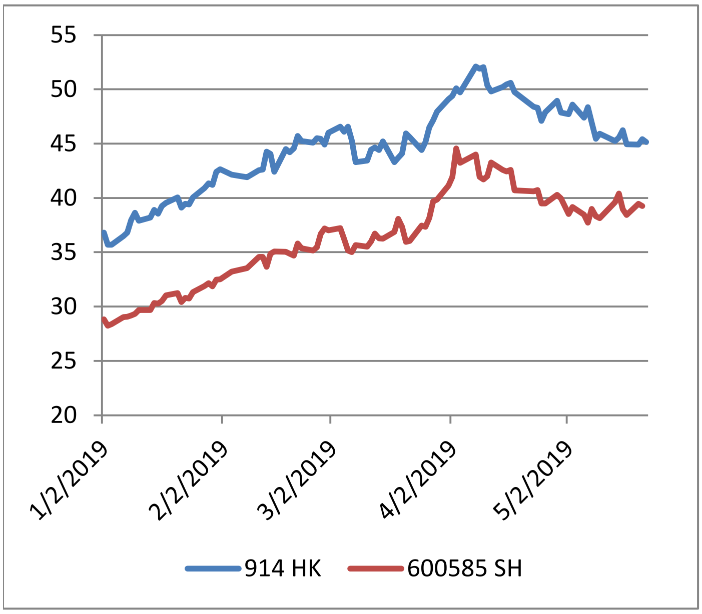
8.5
8
7.5
7
6.5
6
5.5
5
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
390 HK
601390 SH

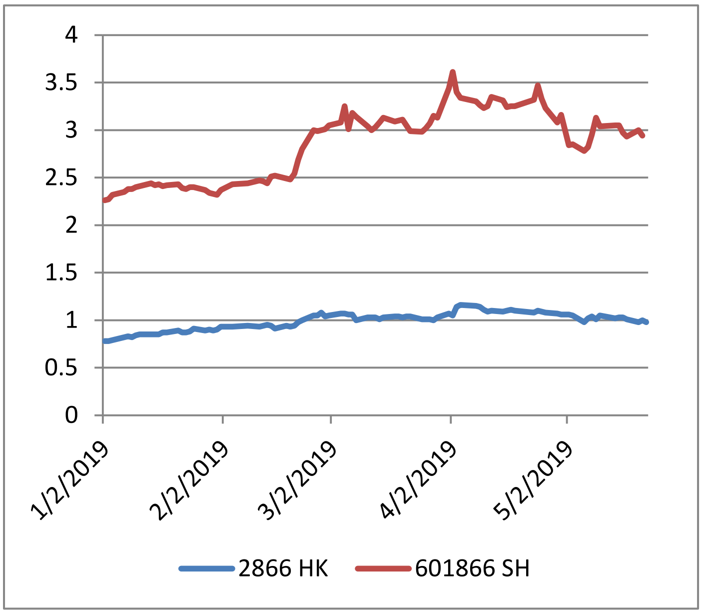


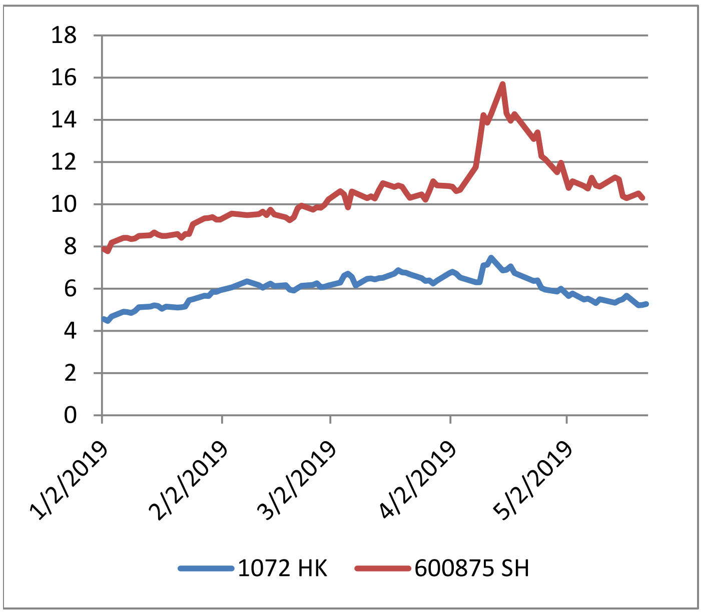
10
9
8
7
6
5
4
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1055 HK
600029 SH

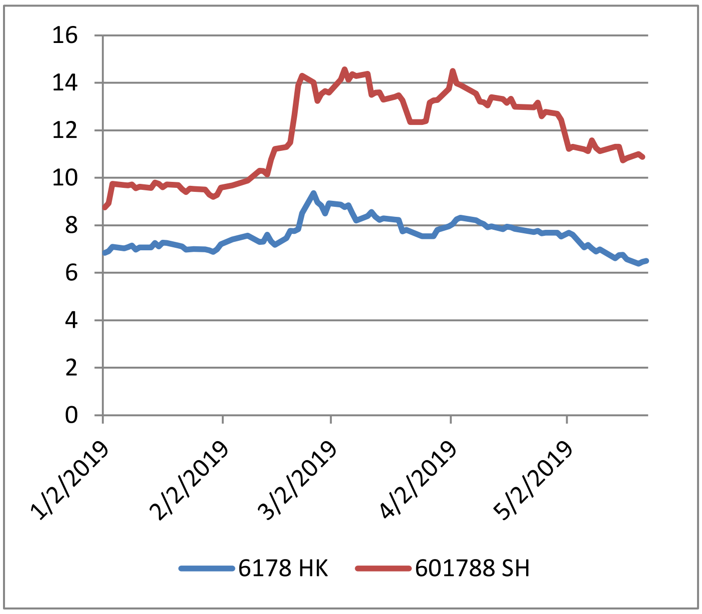


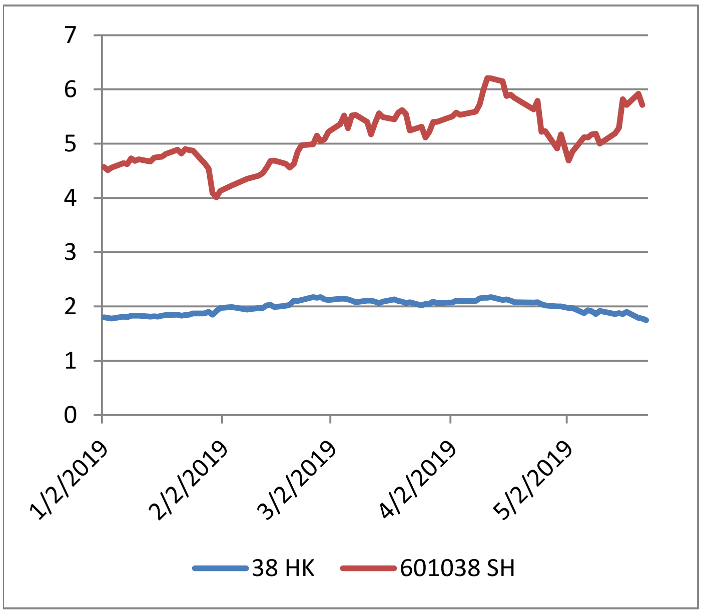
45
40
35
30
25
20
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
3968 HK
600036 SH

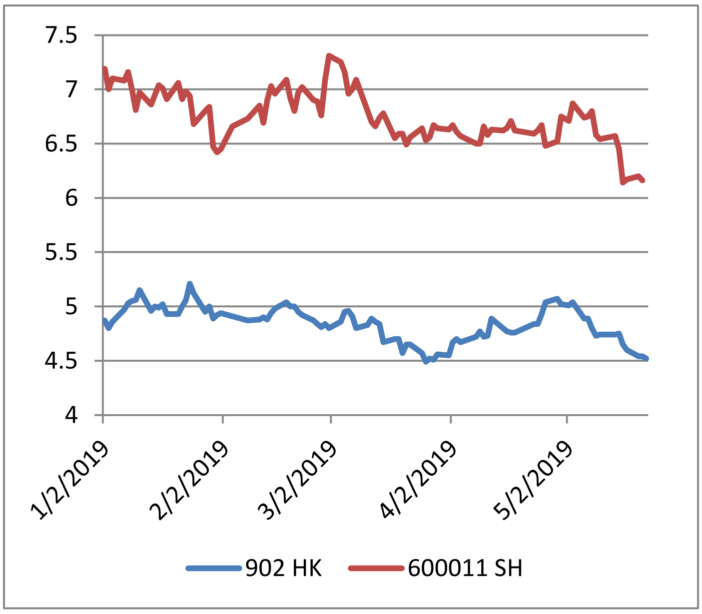

55
50
45
40
35
30
25
20
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
914 HK
600585 SH

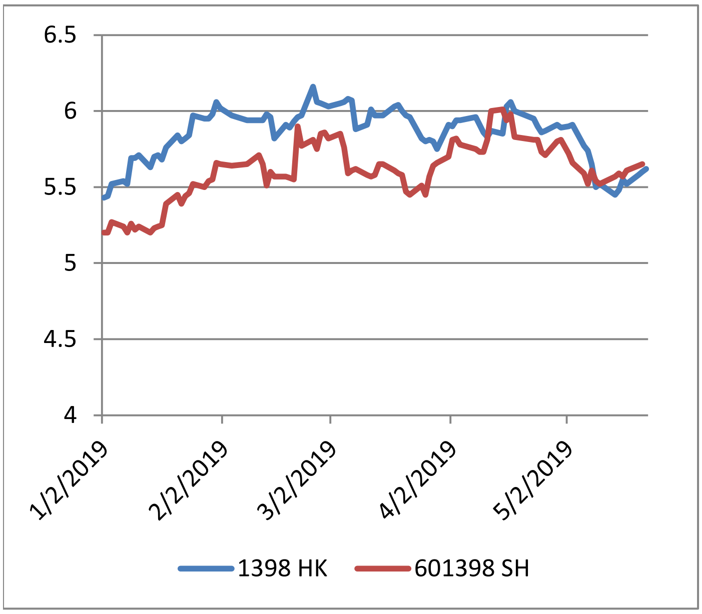

4
3.5
3
2.5
2
1.5
1
0.5
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
2866 HK
601866 SH

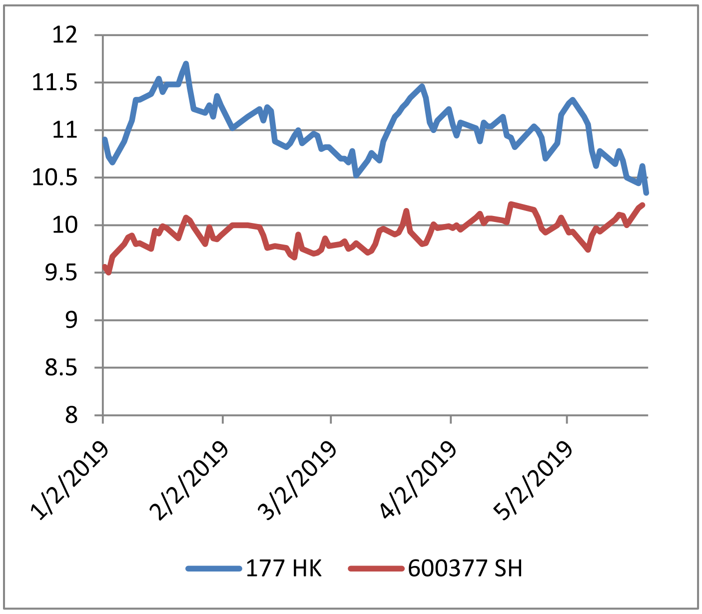

18
16
14
12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1072 HK
600875 SH

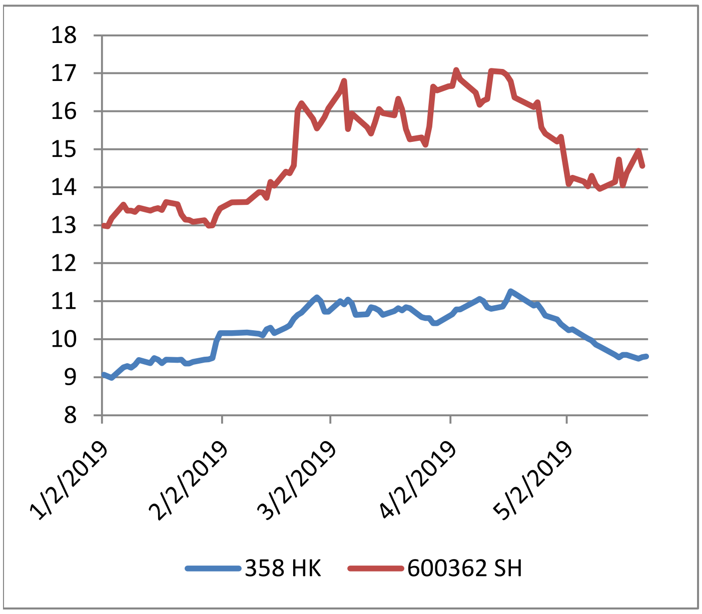

16
14
12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
6178 HK
601788 SH

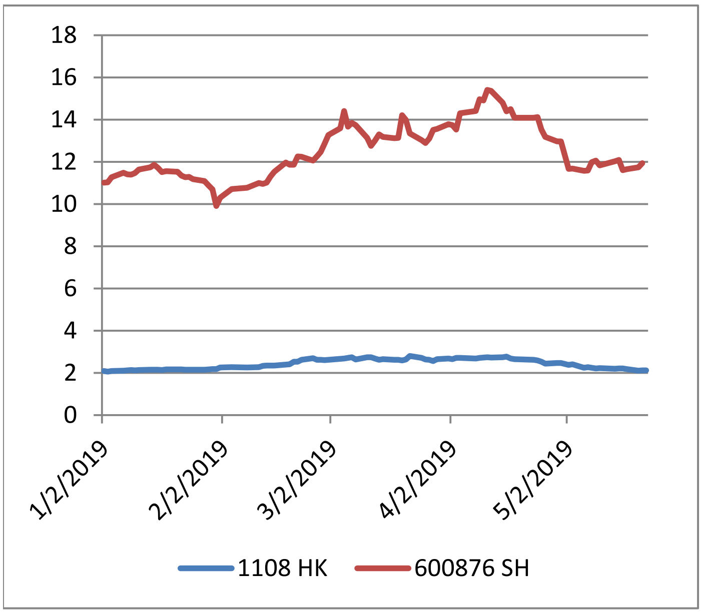

7
6
5
4
3
2
1
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
38 HK
601038 SH

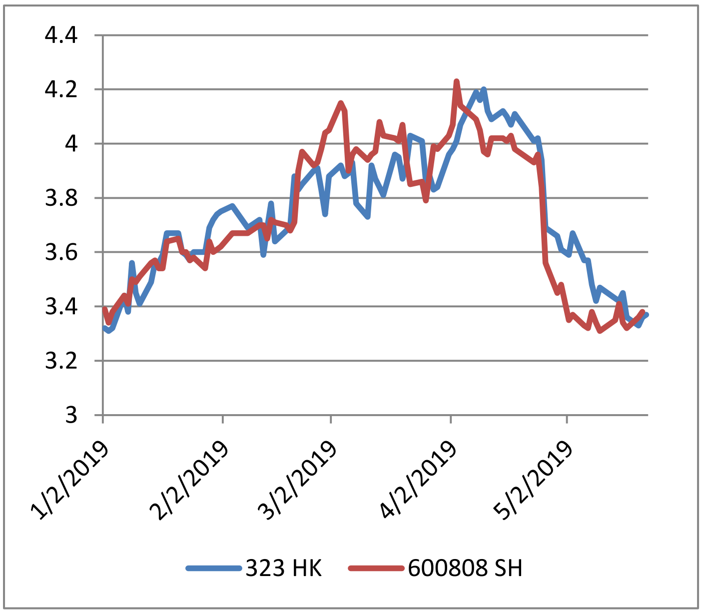

7.5
7
6.5
6
5.5
5
4.5
4
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
902 HK
600011 SH

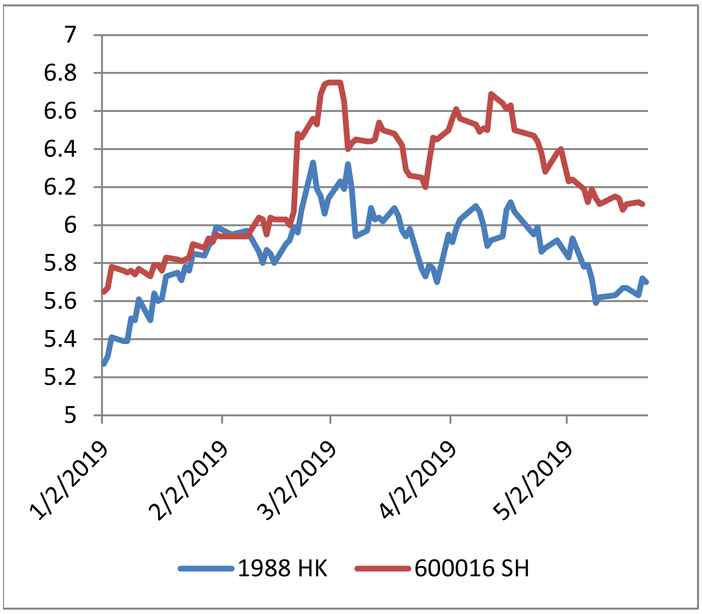

6.5
6
5.5
5
4.5
4
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1398 HK
601398 SH

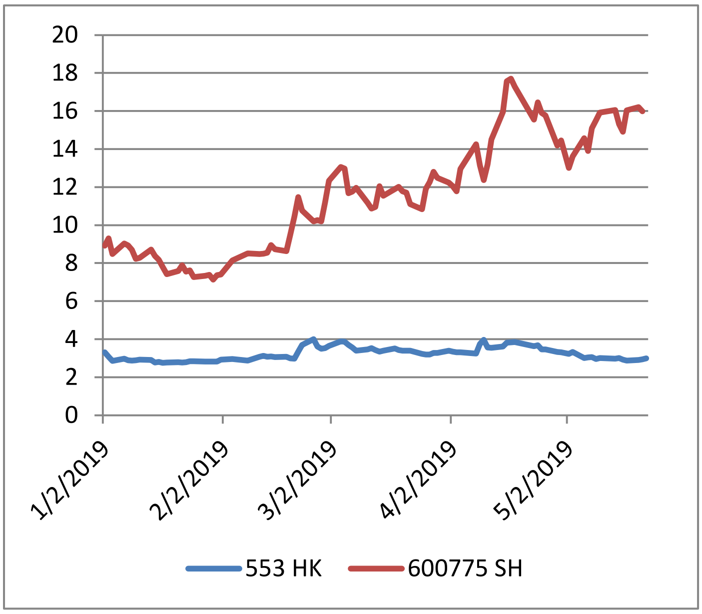

12
11.5
11
10.5
10
9.5
9
8.5
8
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
177 HK
600377 SH

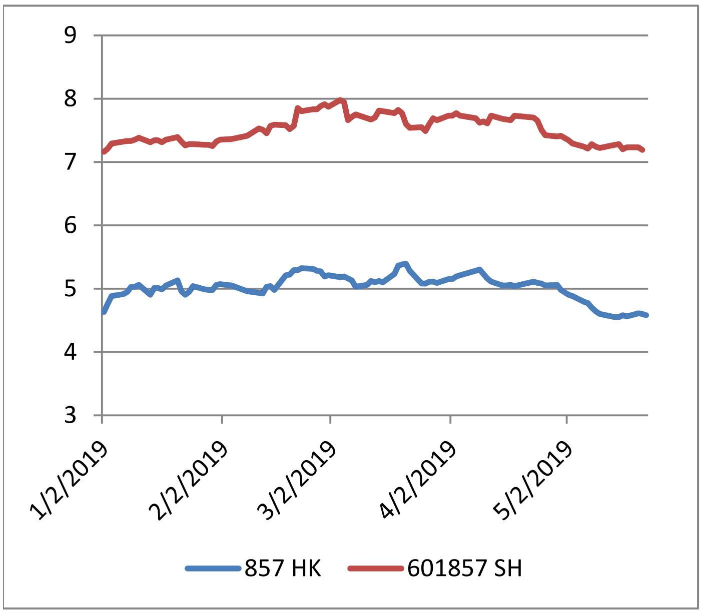

18
17
16
15
14
13
12
11
10
9
8
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
358 HK
600362 SH

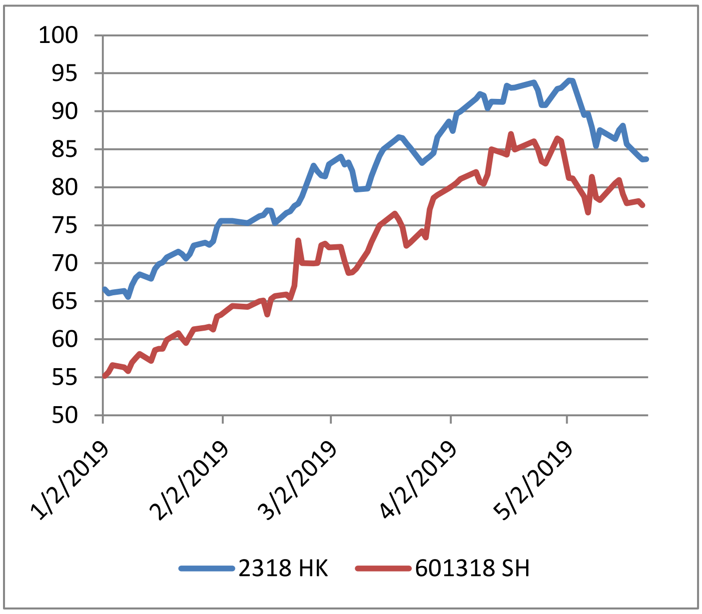

18
16
14
12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1108 HK
600876 SH

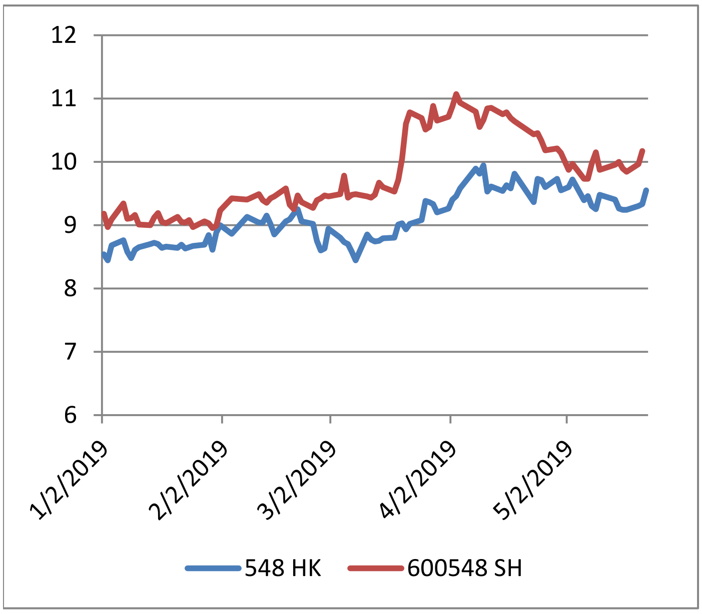

4.4
4.2
4
3.8
3.6
3.4
3.2
3
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
323 HK
600808 SH

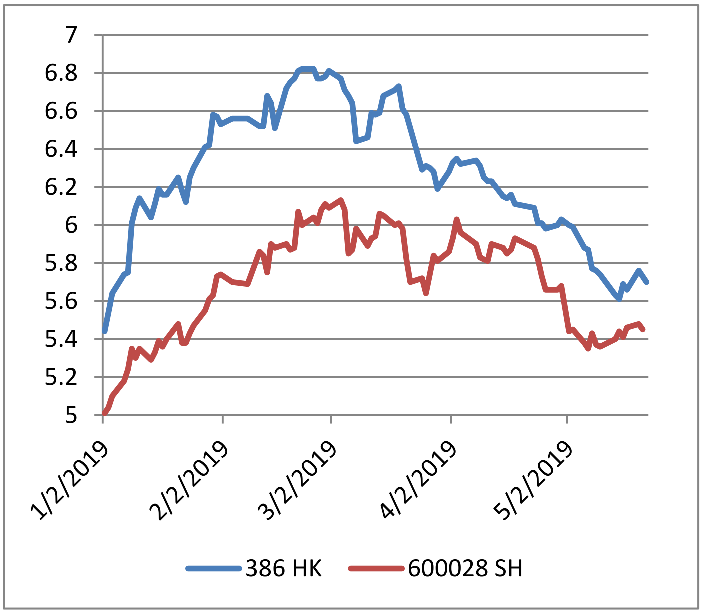

7
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
5
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1988 HK
600016 SH

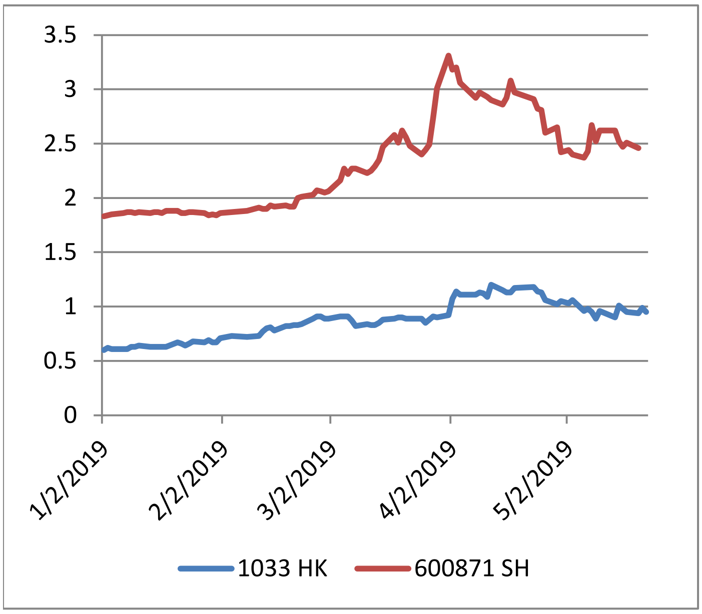

20
18
16
14
12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
553 HK
600775 SH

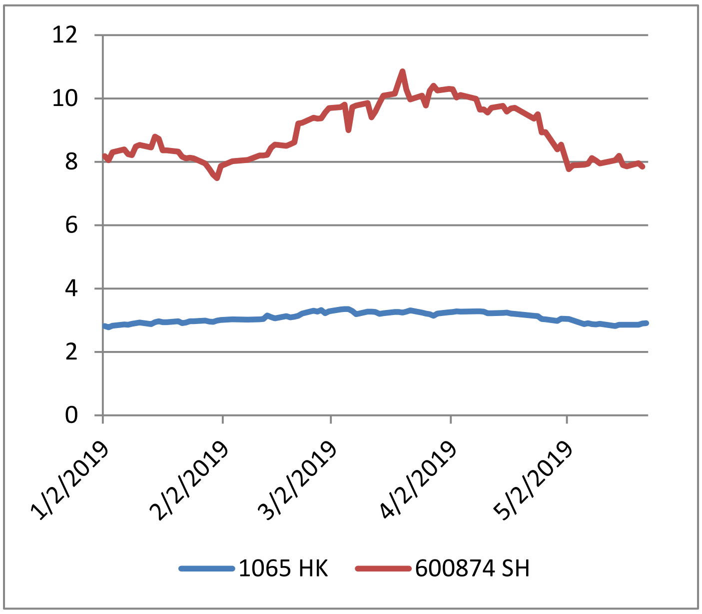

9
8
7
6
5
4
3
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
857 HK
601857 SH

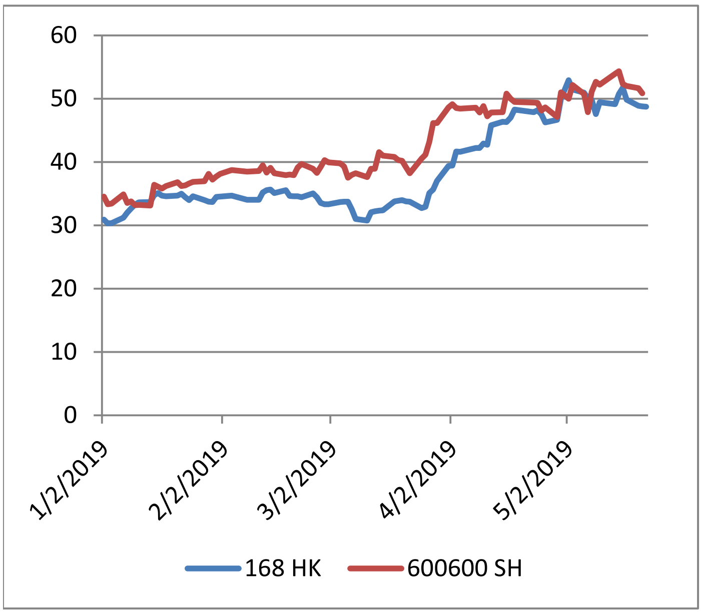

100
95
90
85
80
75
70
65
60
55
50
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
2318 HK
601318 SH

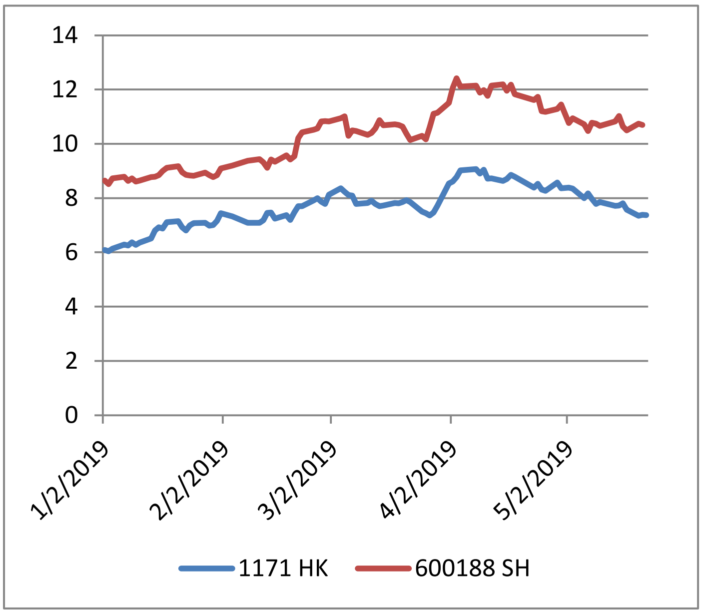

12
11
10
9
8
7
6
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
548 HK
600548 SH


7
6.8
6.6
6.4
6.2
6
5.8
5.6
5.4
5.2
5
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
386 HK
600028 SH


3.5
3
2.5
2
1.5
1
0.5
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1033 HK
600871 SH


12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1065 HK
600874 SH


60
50
40
30
20
10
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
168 HK
600600 SH


14
12
10
8
6
4
2
0
1/2/2019
2/2/2019
3/2/2019
4/2/2019
5/2/2019
1171 HK
600188 SH