# Price discovery and information in an emerging market: Evidence from China*


Jingyun Ma†
Tsinghua University

Peter L. Swan‡
University of New South Wales

Fengming Song§
Tsinghua University


Draft: December 10, 2008


## Abstract

We examine the price discovery process for cross-listing A- and H- shares, the Chinese stocks simultaneously listed in the Mainland China (MC) and the Hong Kong (HK) stock markets. The empirical results of a sample of 30 A-H stock pairs from 2001 to 2007 show that the prices for cross-listing A- and H- shares in MC and HK stock markets are becoming more and more cointegrated, displaying an evolution of the emerging stock market in MC. The MC market contributes to most of the price discovery, according to Hasbrouck's (1995) IS method and the Gonzalo and Granger (1995) PT method. We then combine the results of econometrics method (IS) with the decomposition of the bid-ask spread in market microstructure research. The regression results show a significantly positive relationship between the relative IS of Mainland China and its relative Adverse Selection Component of the effective relative spread, indicating that the reason for MC dominating price discovery is the information advantage of MC domestic investors. The IS of MC relative to HK is also positively related with the Shenzhen Stock Exchange dummy and the 2006 year dummy, when the QDII policy was carried out.



*Key words*: Price discovery, Cross-listing, Cointegration, Chinese securities
JEL Classification: D82, G14, G28, G32, O16


---


*We wish to thank SIRCA (Securities Industry Research Centre of Asia-Pacific) for the provision of Reuters data, the Australian Government for the provision of the Endeavour Research Fellowship to Ma and the Australian Research Council (ARC) for support for Swan. This paper has been accepted for the Third CAF-FIC-SIFR Emerging Market Conference in a very competitive process.



†Doctoral candidate in Finance, School of Economics and Management, Tsinghua University, E-mail: majy.05@sem.tsinghua.edu.cn . Her research was carried out whilst a visiting scholar in the Australian School of Business, UNSW.

‡Banking and Finance, Australian School of Business, The University of New South Wales, Sydney, NSW 2052 Australia. E-mail: peter.swan@unsw.edu.au

§ Professor of Finance, School of Economics and Management, Tsinghua University, E-mail: songfm@sem.tsinghua.edu.cn

## 1. Introduction

One of the most important functions of the security market is price discovery – the impounding and reflecting of new information into the equilibrium security prices that are observable to the market. For the same, or intrinsically related, securities traded at multiple markets, the central concern is how the informational price discovery among these trading venues occurs. In particular, who are the most informed traders and what is the nature and source of their informational advantage? Once we know which venue dominates the price discovery process, it is possible to infer answers to some of these more fundamental questions.

The purpose of this paper is to analyze the price discovery process of the Chinese A- and H- shares, which are stocks issued by the same Chinese companies but traded respectively in the Mainland China (either Shanghai Stock Exchange or Shenzhen Stock Exchange) and Hong Kong stock markets. We show that the financial development of Mainland China's emerging market has a significant impact on the price discovery process. Moreover, using a sizeable period of intraday data (year by year for seven consecutive years), our results indicate an important time-varying characteristic of the cointegration between A- and H- share prices. The empirical results suggests that the Chinese mainland market contributed more to the price discovery process than the Hong Kong market for most of the A-H pair stocks, in most years. Our findings thus support the home-bias hypothesis for cross-listing stock trading. In other words, domestic investors are better informed about domestic stocks and thus price discovery will take place in the domestic market.

In addition, we find that the adverse selection component of bid-ask spread in the A-shares market is significantly positively related with its contribution to price discovery relative to the H-shares market, implying an information advantage for domestic investors in the Mainland China market. These findings are supportive of the hypothesis that the well-established home bias of investors towards domestic stocks is due to geography and proximity. Traders with the greatest contact with and proximity to informed insiders have a powerful informational advantage that is reflected in far stronger price discovery in the local market.

In previous research, several econometric approaches have been applied to intraday data of securities simultaneously traded at different security exchanges (even in different countries), in order to measure the contributions of price discovery for each trading venue. The literature can be mainly divided to two categories according to their econometric methods and measures of contribution, respectively the Hasbrouck (1995) Information Share (IS) method and the Gonzalo and Granger (1995) Permanent-Transitory decomposition (PT) method.

The first category is the IS method initiated by Hasbrouck (1995), which assumes an implicit unobservable efficient price common to all markets and defines the information share of a particular market as the proportional contribution of that market's innovations to the innovation in the common efficient price, building on his earlier contributions (Hasbrouck 1991a, b). The applications or adjusted applications of this method can be found in Booth et al. (2002), Chan et al. (2002), Hupperets and Menkveld (2002), Eun and Sabherwal (2003), Hasbrouck (2003) and Chakravarty et al. (2004), to name but a few.

The second category is the common factor method characterized by Gonzalo and Granger (1995)'s PT decomposition of cointegrated time series, applied in Harris et al. (1995, 2002a), Booth et al. (1999), Booth et al. (2002) and Chakravarty et al. (2004). It defines each market's contribution to price discovery as a proportional function of its error correction adjustment coefficients.

Both the IS and the PT methods are based on a vector error correction model (VECM), but using different definitions of contribution to price discovery and decomposition techniques. A special issue of *Journal of Financial Market* has a profound discussion on the similarity and differences between both methods, including the arguments by the initiators. Detailed econometric analysis, comparison of simulation results and theoretical explanation in this special issue include the papers of de Jong (2002), Baillie et al. (2002), Lehmann (2002), Harris et al. (2002b), and Hasbrouck (2002). The main findings are that the two models provide similar results if the residuals are uncorrelated between markets, but typically different results when substantial correlation exists.

Other than the above econometric approaches which seek to provide an explicit measure of each market's contribution to price discovery, there are also approaches that examine the process of information being incorporated into security prices, by using bid-ask spread decompositions (e.g., Lin et al. (1995), Huang and Stoll (1997)), structural model of intraday price formation (e.g., Madhavan et al. (1997)) and microstructure effects on price, such as impacts of block trading (e.g., Kraus and Stoll (1972)).

When these econometric and microstructure methods have been applied to analyze price discovery for cross-listing stocks in different countries, the conclusions vary for different regions and stock exchanges. We make a concise summary of the literature on price discovery for stocks cross-listed in different countries in Table 1, presenting their cross-listing countries, methodologies and main results. Most of these studies focus on the stock markets for two different countries, typically one of which is the U.S., and investigate the mutually adjusting pricing process of the dually listed stocks. According to which market, either home market or foreign market, dominates the price discovery process, the literature can be divided to the following categories. We provide details of this literature in Table 1.

(Insert Table 1 about here)

(1) The first category indicates that the dominant price discovery location varies cross-sectionally for different stocks (domination varies). Werner and Kleidon (1996) analyze intraday patterns for U.K. and U.S. trading of British cross-listed stocks and the two-hour overlap is characterized by concentrated trading, as private information originating in New York gets incorporated into prices in both markets. Hupperets and Menkveld (2002) study for Dutch stocks cross-listed at NYSE during the one overlapping trading hour and suggest no consistent price discovery origin across all stocks, either Amsterdam or New York. Eun and Sabherwal (2003) examine the U.S.-listed Canadian stocks and find the U.S. of price discovery ranging from a percentage of 0.2 to 98.2, with an average of 38.1, which is directly related to the U.S. share of trading, informative trades and inversely related to U.S. bid-ask spreads. Wang and Jiang (2003) use daily data to find that H-shares exhibit significant exposure to Hong Kong market factors and behave more like Hong Kong stocks than mainland Chinese stocks, but the data frequency is too coarse. Halling et al. (2007) investigate the stock trading for firms with a U.S. cross-listing and present several reasons for companies with U.S. large trading volumes, respectively for companies from emerging countries and developed countries.

(2) The second category (home dominates) explicitly presents empirical results of home-market dominated price discovery process. Biais and Martinez (2004) study opening prices set simultaneously for German-French dual-listing stocks in Frankfurt and Paris, to find that home prices are informationally more efficient than foreign prices. Grammig, Melvin and Schlag (2005) analyze exchange rates along with equity quotes for three German firms listed in New York (NYSE) and Frankfurt (XETRA) during overlapping trading hours. They find that most (but not all) of the fundamental or random walk component of firm value is determined in Frankfurt. Pascual, Fuster and Climent (2006) introduce a methodology that distinguishes trade-related and trade-unrelated informative shocks. They also provide empirical evidence that NYSE has made almost no contribution to price discovery of Spanish cross-listed stocks during the daily (two-hour) overlapping interval. Menkveld, Koopman and Lucas (2007) use a state space model to study 24-hour (round-the-clock) price discovery for Dutch stocks cross-listed in the U.S. market and suggest a minor role for the NYSE in price discovery for Dutch shares. This is in spite of its non-trivial and growing market share.

There have been a few studies on the price discovery for cross-listing stocks of companies in emerging market countries. For example, Domowitz, Glen and Madhavan (1998) use the example of Mexican stock exchange to show that the impact of cross-listing is complex: balancing the costs of order flow migration against the benefits of increased intermarket competition. However, they do not analyze which market dominates the price discovery.

From a broad perspective, most of these articles merely truncate a relatively short period of time and investigate price discovery during this one-off period. What is neglected from their research, but is of great importance, are the changes taking place over time. Therefore, we use an intraday sample covering seven years and describe the time-varying trend by studying year by year and their series. We use the Johansen (1991) cointegration test, the Hasbrouck (1995) IS method and the Gonzalo-Granger (1995) PT decomposition method to analyze each cointegrated year for each A-H pairs stock. We use a unique tick-by-tick data set to study the rapidly-developing period of the Chinese emerging stock market, compared with the developed Hong Kong market. Our results provide strong evidence of the emerging market's evolution in terms of its contribution to price discovery of the cross-listing A- and H- shares.

We then combine our econometric methods with an important bid-ask spread decomposition from the market microstructure literature, merging up until now distinct literatures, so as to investigate the driving forces behind our empirical results.

This paper is organized as follows: Section 2 describes the model setting and econometric methods used – Johansen's (1991) cointegration test, Hasbrouck's (1995) IS, and Gonzalo-Granger's (1995) PT decomposition. In Section 3, we introduce the trading mechanism and intraday data of our Mainland China A-shares and Hong Kong H-shares markets. Detailed empirical applications of the econometric methods and analysis of results are presented in Section 4. In Section 5, we decompose the bid-ask spread to explain the Hasbrouck (1995) IS in terms of fundamentals such as the decomposed components of bid-ask spread and other control dummy variables. Section 6 concludes with a brief summary of our study and results.

## 2. Model and Methods

In this section, we demonstrate the cointegration model and error correction models being used. We show the differences between the IS method and PT method, particularly the difference in their interpretation of each market's contribution to price discovery. Consider a pair of stocks of the same company trading at two different venues. Since they are assets which share a common future cash flow and other common company features, such as risks, we expect there to be a strong arbitrage force closely linking their prices[1]. As a result, we model their prices to discover the common long-term component which plays the binding role between prices.

### 2.1 Cointegration model

Cointegration is a linear dynamic model used to describe the long-run relationship between non-stationary time series data. Let the (log) price vector $p_t = (p_{ct}, p_{ht})'$ denote prices at time t for two markets, where $p_{ct}$ is the price for Mainland China stock market at time t and $p_{ht}$ the price for Hong Kong stock market at time t. In our market microstructure applications, we assume the stock prices follow a vector

[1] Although there is large price discount between Chinese A- and H- shares, our cointegration test show that for certain time periods for these stocks, the arbitrage force links their prices in spite of transaction costs and other impediments. Chen and Swan (2007) explain the sizeable price differences between otherwise identical Chinese A- and B- shares in terms of a liquidity asset pricing model.

autoregressive process and contain an independently and identically distributed random walk component which represents the random arrival of information. As a result, each price series of two stock markets are non-stationary and integrated of order 1, as denoted $I(1)$. The stock returns $\Delta p_t$ (log price changes) are assumed to be a covariance stationary time series process of $I(0)$. The implied vector moving average (VMA) form of the prices can be expressed as:

$$\Delta p_t = \Psi(L)\varepsilon_t = \varepsilon_t + \Psi_1\varepsilon_{t-1} + \Psi_2\varepsilon_{t-2} + \cdots, \qquad (1)$$

where $\varepsilon_t$ is a zero-mean $2\times1$ vector of serially uncorrelated disturbances with covariance matrix of $\Omega$, satisfying:

$$E[\varepsilon_t\varepsilon'_s] = \begin{cases} 0 & if\ t \neq s; \\ \Omega & otherwise \end{cases},$$

$L$ is the lag operator and $\Psi$ is a polynomial in the lag operator. Since the price differences are stationary, in this case, the prices vector process is cointegrated of order 1, as $CI(1\ ,\ 1)$ with a cointegrating vector $\beta = (1,-1)'$, satisfying the condition that $\beta' p_t$ is $I(0)$.

This implies that $\beta'\Psi(1) = 0$, where $\Psi(1) = I + \Psi_1 + \Psi_2 + \cdots$, is the sum of the moving average coefficients and has rank 1. According to our assumption regarding $\beta$, the two rows of $\Psi(1)$ are identical and the long-run impacts of an innovation $\varepsilon_t$ on each of the prices are the same, represented by $\Psi(1)\varepsilon_t$. In our two market case, denote $\psi = (\psi_c, \psi_h)'$ as the common row vector of $\Psi(1)$ and, as suggested by Stock and Watson (1988), the cointegrated price process has the following common trends representation:

$$p_t = p_0 + \psi(\sum_{s=1}^{t}\varepsilon_s) + \Psi^*(L)\varepsilon_t. \qquad (2)$$

This representation of $p_t$ consists of three parts: a 2-vector of initial price values which reflect the non-stochastic differences between two prices, a random walk component fundamentally common to two prices changes, and a zero-mean covariance stationary process.

We use the Johansen (1991) cointegration test in Section 4 to estimate and test for the cointegration relationship between pairs of A- and H-shares prices year by year from 2001 to 2007.

**2.2 Information Share (IS) method**

According to the price representation of (2), when new information arrives, the increment $\psi\varepsilon_t$ is the component of permanently impounded price change, with its variance $\psi\Omega\psi'$. If $\Omega$ is diagonal, which implies that the innovation terms of each market are not correlated, we can determine each market's contribution to the innovation of the random-walk component $\psi\varepsilon_t$. Hence, the IS of each market $j$ ($c$ for Mainland China and $h$ for Hong Kong) is defined as the proportion of its innovation to the total variance:

$$IS_j = \frac{\psi_j^2 \Omega_{jj}}{\psi\Omega\psi'} = \frac{\psi_j^2 \omega_j^2}{\psi_c^2 \omega_c^2 + \psi_h^2 \omega_h^2}, \qquad j = c, h, \tag{3}$$

where $\psi_j$ is the $j^{th}$ element of $\psi$ and $\omega_j^2$ the $j^{th}$ diagonal element in $\Omega$. By normalization, the sum of the information share $IS_c + IS_h$ is equal to 1, and a higher (lower) IS for market $j$ indicates more (less) incorporation of innovation from new arrival of information.

However, the above representation comes from the assumption that there is no correlation between the price innovations across two markets. If the covariance matrix $\Omega$ is non-diagonal, we follow Hasbrouck (1995) and use the Cholesky factorization to triangularize the covariance matrix $\Omega$. Letting $F$ denote the lower triangular Cholesky factorization of $\Omega$ such that $FF' = \Omega$, the IS of $j^{th}$ ($c$ for Mainland China and $h$ for Hong Kong) market can be redefined as

$$IS_j = \frac{([\psi F]_j)^2}{\psi\Omega\psi'}, \qquad j = c, h, \tag{4}$$

where $[\psi F]_j$ is the $j^{th}$ row of the matrix $\psi F$. Since the factorization imposes a maximized (minimized) information share on the first (last) price, an upper (lower) bound of $j^{th}$ market's information share can be calculated by permuting $\psi$ and $\Omega$ to place that market's price the first (last) one. In our two market case, we obtain the upper and lower bounds for Mainland China and Hong Kong market respectively and report the mid-point of the upper and lower boundaries of each market's IS.

**2.3 Permanent-Transitory (PT) method**

According to Gonzalo and Granger (1995), since both price series are integrated of order 1, $I(1)$, they can be decomposed to the following PT form:

$$p_t = \Gamma f_t + A z_t = \Gamma \gamma'_{\perp} p_t + A\alpha' p_{t-1}, \tag{5}$$

where $f_t$ represents the permanent common component and $z_t$ the transitory component. $\Gamma$ and $A$ are respectively the loading matrix for permanent component coefficient and transitory component. By construction, the permanent factor $f_t$ is $I(1)$ and the error correction term $z_t$ is stationary, as $I(0)$. The error correction term does not Granger-cause the permanent factor in the long run. Gonzalo and Granger (1995) show that under these assumptions and restrictions, the matrices $\Gamma = \alpha_{\perp}(\gamma'_{\perp}\alpha_{\perp})^{-1}$ and $A = \gamma(\alpha'\gamma)^{-1}$ , where $\gamma = (\alpha'_{\perp}\beta_{\perp})^{-1}\alpha'_{\perp}$ . By definition, the vector of common permanent factor weights $\gamma_{\perp}$ is orthogonal to the coefficient vector $\gamma$ such that $\gamma'_{\perp}\gamma = 0$.

Since we construct the error correction term such that $\beta = (1,-1)'$, the permanent component $\Gamma f_t$ is a weighted average of observed prices. Therefore, the PT decomposition suggests using each price's weight of this permanent component as its share of price discovery:

$$PT_j = \frac{\alpha_{\perp,j}}{\alpha_{\perp,c} + \alpha_{\perp,h}}, \quad j = c, h. \tag{6}$$

In addition, since we have $\alpha'_{\perp}\alpha = 0$, the coefficient of the error correction term $\alpha = (\alpha_c, \alpha_h)'$ can be used to express the price discovery share in (6), as

$$PT_c = \frac{\alpha_h}{\alpha_h - \alpha_c}, \; PT_h = \frac{-\alpha_c}{\alpha_h - \alpha_c}.$$

**2.4 Comparison of the IS and PT methods**

It is important to point out that although both IS and PT methods try to measure the contribution to price discovery for cross-listing securities, their results may not be exactly the same, because their definition of share of price discovery vary from each other. A detailed discussion of these methods can be referred to in the special issue of *Journal of Financial Market* in 2002, as already noted. Neither of these two methods attempts to determine which market has the "best" price. They try to use different measurements to investigate which market moves first in reaction to new information, and which market is more efficient in incorporating new information. The IS method emphasizes each market's contribution to innovation of new information and the PT method emphasizes the adjustment of each market price to their long-run cointegrated relationship.

## 3. Data Description

In this section, we briefly introduce the history and relationship between stock exchanges of Mainland China and Hong Kong. We also compare the trading hours and features of both markets. A summary statistics of the cross-listing companies is then presented and details of the sample intraday data are discussed.

### 3.1 Two markets and cross-listing

The price discovery process analyzed in this paper involves stocks of 30 Chinese companies cross-listed in the Mainland China and Hong Kong stock markets, respectively as A- and H- shares. There are two stock exchanges in Mainland China: the Shanghai Stock Exchange (SSE) established in November 1990 and the Shenzhen Stock Exchange (SZSE) established in July 1991. While in Hong Kong (Special Administrative Region of China), in March 1999, under the comprehensive market reform of the stock and futures markets, the Stock Exchange of Hong Kong Limited (SEHK), Hong Kong Futures Exchange Limited (HKFE) and Hong Kong Securities Clearing Company Limited (HKSCC) merged and became wholly-owned subsidiaries of a single holding company, HKEx. Companies incorporated in China can issue A-shares to its domestic investors and traded in Chinese currency Yuan, listed either in SSE or in SZSE. Under the approval of the China Securities Regulatory Commission (CSRC), these companies can also issues H-shares to international investors in Hong Kong stock market, traded in Hong Kong dollars and listed in HKEx.

On 19 June 1993, the SEHK, the CSRC, the SSE and the SZSE signed a Memorandum of Regulatory Co-operation in Beijing, which paved the way for the listing of mainland-registered enterprises in Hong Kong. On 15 July 1993, Tsingtao

Brewery Company Limited became the first mainland enterprise to list its H-shares on the Stock Exchange. Most of the cross-listing companies firstly issued H-shares in Hong Kong and return to Mainland China to issue A-shares and become cross-listing stocks in these two markets. In recent years, there have been more and more companies simultaneously cross-listed in Mainland China and Hong Kong. Until June 3, 2008, this number counts to 60 pairs of cross-listing A- and H- shares. The reason why we only choose 30 cross-listing pairs is that all these 30 pairs cross-listed before or in the year 2006 so that we have at least two years (our sample covers intraday data from the year 2001 to 2007) to analyze. Other cross-listing stocks are eliminated from our sample either because of their relatively short listing period or incompleteness in intraday data.

The opening hours and trading mechanisms are the same for SSE and SZSE, but differ between SSE (SZSE) and HKEx. They all conduct trading from Monday to Friday except for their public holidays, which could differ for two regions. For each trading day, from Figure 1, we can clearly see the overlapping and non-overlapping trading hours for the Mainland China stock market and the Hong Kong stock market. There are three trading periods for both markets: the pre-open period with call auction, the morning session with continuous auction, and after a lunch break, the afternoon session with continuous auction. For our purpose, we choose the overlapping periods of each morning session and afternoon session for both markets with continuous auctions. Those periods are: from 10:00 ~ 11:30 in the morning and from 14:30 ~ 15:00 in the afternoon.

(Insert Figure 1 about here)

In addition, our sample eliminated the days when only one market has data, since our cointegration and price discovery process analysis requires the interaction between two markets.

### 3.2 Data sources and sample details

We construct our sample of A- and H- shares intraday data from SIRCA (Securities Industry Research Centre of Asia-Pacific) database. It provides the quote and transaction intraday data for stocks traded in Mainland China and Hong Kong stock markets, sorted by symbol (company name). As we have explained, we choose 30 A- and H- shares pairs which cross-listed before or in 2006 and the data of which are complete in this database. In order to present the time-varying characteristic of our result and the development of the emerging market, our intraday sample covers seven years, from 2001 to 2007. Table 2 shows the pairs of cross-listing stocks, with the company names, A- and H- share codes and their date of IPOs on each market:

(Insert Table 2 about here)

Since most stocks do not have intraday data in Mainland China until 2003 for technological reasons, we can only cover the earliest possible date for them during the period from 2001 to 2007. Our final sample consists of 26 companies cross-listed in SSE and HKEx, and four companies cross-listed in SZSE and HKEx. From first-listing locations, we can clearly see that most of the sample companies (28 out of 30) firstly made IPOs in Hong Kong stock market, only ZTE and CHINA MERCHANT BANK firstly made IPOs in Mainland China. This is directly related to the policy direction of the Chinese government in the 1990s.

The analysis of cross-listing stocks can be based on either transaction prices or quoted prices. Since the use of transaction prices could suffer from bid-ask bounce, we adopt the mid-quote price sampling method to perform our analysis. In particular, we select the midpoint of the bid and ask quotes in each market at 1-minute fixed time interval.[1] If there are no quotes at exactly the 1-minute time point, we use the bid-ask quotes for the closest happened transaction to each time point. Consequently, for each stock in each market, we have 122 prices for analysis per day (10:00 ~ 11:30 in the morning and 14:30 ~ 15:00 in the afternoon). Since A-shares in SSE (SZSE) are traded in Chinese *Yuan* and H-shares in HKEx are traded in *Hong Kong dollars,* we use the corresponding days' exchange rates between *Yuan* and *Hong Kong dollars* to transfer the prices of H-shares to comparable Chinese *Yuan* to make the prices consistent in terms of denomination and magnitude. We obtain the everyday exchange rate between *Chinese Yuan* and *Hong Kong dollars* from the Datastream database.

## 4. Empirical Results and Analysis

### 4.1 Results of cointegration test

[1] For robustness and consistency, we also perform the analysis by using 5-minute and 10-minute intervals and there is no qualitative change in results.

As described in the data section, we use the 1- minute-interval, mid-quote time-series intraday prices for pairs of A-share and H-share in each year to estimate the Johansen (1991) cointegration test statistics. The estimations and results are presented in Table 3, with the asterisk implying a cointegration relationship at a 5% significant level between A- and H- shares prices for a certain stock pair in a certain year. The last four rows in Table 3 respectively shows the number of cointegrated A-H pairs, the number of total A-H pairs, the frequency of significant A-H cointegrated pairs and the percentage of significant A-H cointegrated pairs in each year from 2001 to 2007.

(Insert Table 3 about here)

We also draw the line graph for the absolute number and percentage of cointegrated A-H pairs in each year. From Figure 2, we can clearly see the upward trend of cointegrated A-H pairs from 2001 to 2007. Although the percentage of cointegrated A-H pairs went up to 50% in 2002, the total sample of 4 pairs claims that this point does not mean a higher cointegrated level than the year 2003 or 2004. Therefore, we can conclude that the prices of A-shares market of Mainland China and the H-shares market of Hong Kong are becoming more and more cointegrated from 2001 to 2007. This can be attributed to several reasons, such as the rapid development of China’s financial market, the gradual floating policy of Chinese currency Yuan respective to American dollars and the gradual opening up of China’s domestic capital market, the most important of which is the QFII and QDII policy adopted by Mainland China in this period.

(Insert Figure 2 about here)

Qualified Foreign Institutional Investor (QFII) is a scheme which allows qualified foreign financial institutions to trade Chinese A-shares via special accounts opened at designated custodian banks. Similarly, Qualified Domestic Institutional Investor (QDII) is a capital market scheme set up to allow financial institutions in Mainland China to invest in offshore markets such as securities and bonds in Hong Kong. They are both transitional arrangements which provide limited opportunities for domestic (foreign) investors to access foreign (domestic) markets at a stage where a country/territory’s currency is not traded or freely floated and where capital is not able to be transferred freely in and out of the country. On November 5, 2002, the China Securities Regulatory Commission (CSRC) and the People's Bank of China (PBoC) introduced the QFII program to provide a means for foreign capital to access China's financial markets. On 13 April 2006, the Chinese government announced the QDII scheme, allowing Chinese institutions and residents to entrust Chinese commercial banks to invest in financial products overseas, limited to fixed-income and money market products. Furthermore, the Chinese government announced on 11 May 2007 a scheme to widen the scope of the QDII investment so that banks can invest in stocks-related products, but with certain restrictions. This deregulation of investment between Mainland China and Hong Kong since 2002 provides more opportunities for financial institutions to trade simultaneously in both markets and to seek arbitrage opportunities. Especially after the introduction of a policy to allow Mainland banks to invest in stocks in Hong Kong market in 2007, we observe a rapid increase in cointegrated A-H pairs.

Next, we report the empirical results for IS method and PT method. Since both models base their analysis on the vector error correction model (VECM) explained previously, which requires the two time-series of intraday prices to be cointegrated, we select just the cointegrated years for these 30 stocks and apply both methods to the data subset.

**4.2 Empirical results of Information Share (IS) method**

According to the Cholesky factorization described in the method section, we obtain the upper and lower bounds of IS for each stock in each market. Our results show that the upper and lower bounds are not widely separated from each other (within 0.2), which indicates that the correlation of innovations between two markets is not large. Therefore, for simplicity and provide for easy comparisons, we report the mid-point of the upper and lower bounds of IS of Mainland China and Hong Kong stock markets in Table 4. For each stock pair in each cointegrated year, the IS of MC and HK sum up to 1, according to our methodology. Consequently, an IS higher than 0.5 in a certain market indicates that this market’s contribution to innovation is larger than the other market and its contribution to price discovery is larger than the other market, in terms of the IS explanation.

(Insert Table 4 about here)

From Table 4, it is clear that, for almost all pairs of A-H stocks in their cointegrated years, the Mainland China market’s ISs are higher than 0.5. According to our IS model, a higher (lower) IS for a certain market indicates its ability to contribute more (less) to the innovation from new arrival of information. Therefore, our empirical results provide strong evidence that the Mainland China market plays a predominant role in price discovery for cross-listing A- and H- shares. Although the IS for Mainland China market ranges from 0.496 to 0.950 and the IS for Hong Kong market ranges from 0.050 to 0.505, of all the 77 cointegrated pairs used in this sample, only six ISs for Mainland China are higher than 0.7. This indicates that, although Mainland China market dominates the price discovery for cross-listing A- and H- shares, its relative informativeness is only slightly larger than the informativeness of Hong Kong's market.

By contrast, we do find that for Bank of China (BOC), the A-shares of which are listed on July 5, 2006 and the H-shares of which are listed on June 1, 2006, its IS in Hong Kong stock market is slightly larger than that of Mainland China. A possible explanation is that Bank of China in Hong Kong was established and incorporated some years early in 2001. It has been a very important part of BOC, since historically it has been a commercial bank focusing on overseas business. Therefore, it is reasonable to believe that investors in BOC in Hong Kong have a greater information advantage than Mainland China investors.

We present the 77 points of Information Shares for A-H pairs in their cointegrated years in the following scatter graph (Figure 3). Because we take the mid-point of each market's upper and lower Information Shares, the normalized ISs in Mainland China and Hong Kong for each A-H pair in a certain year sum up to 1, as $IS_c + IS_h = 1$. Hence, we can observe a straight line with a slope of -1 in Figure 3. Since most ISs of Mainland China market are higher than those of Hong Kong market, the scattering points mainly occupy the lower right triangle of the square.

A natural question to ask at this stage is: why the Mainland China market has a leading role in price discovery and why the trading in Mainland China is more informative than the Hong Kong market for cross-listing A- and H- shares? We will explore the reasons after presenting the results for PT method in the following section.

(Insert Figure 3 about here)

**4.3 Empirical results of Permanent-Transitory (PT) method**

As we have explained in the model section, the PT decomposition suggests using each price's weight of permanent component as its share of price discovery. We apply this method to the A-H pairs in their cointegrated years and obtain empirical results for shares of price discovery respectively for the Mainland China and Hong Kong markets. Because of normalization, the shares of price discovery for both markets for the same pair in a certain cointegrated year sum up to 1. Therefore, the same as for the IS method, a share of price discovery for a certain market higher than 0.5 indicates a leading role in price discovery for this market. We bold the shares of price discovery for Hong Kong market when they are higher than 0.5, which means that in these years and for these A-H pairs, the Hong Kong market plays a leading role in price adjustment.

(Insert Table 5 about here)

From Table 5, it is not clear which market predominantly leads price discovery and price adjustment. For the A-H pairs in their cointegrated years, 21 out of 77 results imply that the Hong Kong market plays a leading role in price discovery in terms of PT method. For the rest 56 results, the Mainland China market has a share of price discovery higher than 0.5 and leads the price adjustment for cross-listing A- and H-shares. To be clearer, in Figure 4, we present the 77 points of shares of price discovery using PT method for A-H shares in their cointegrated years. Because of normalization, the shares of price discovery in Mainland China and Hong Kong for each A-H pair in a certain year sum up to 1. Hence, we can also observe a line with a slope of -1 in Figure 4 for the scattering points of shares of price discovery. For the lower right triangle of the box, the scattering points falling into this area imply that the Mainland China market has a higher share of price discovery. Correspondingly, the points falling into the upper left triangle of the box show that the Hong Kong market has a higher share of price discovery for certain A-H pairs in these certain years.

(Insert Figure 4 about here)

To illustrate the empirical results from Figure 4 and Table 5, it is not justified to say that the Mainland China market significantly dominates price discovery for all sample pairs in their cointegrated years. However, 72.7% (56/77) of the results do indicate that Mainland China market's share of price discovery is higher than that of the Hong

Kong market and that Mainland China market lead the price adjustment for A-H pairs in these years.

### 4.4 Comparison of the results from two methods

Although the results from the two methods, respectively IS and PT, are not exactly the same as each other, it is reasonable to draw the conclusion that Mainland China market dominates the price discovery for most cross-listing A- and H- shares during their cointegrated years. Notice that in Section 2.4, we have compared the differences between these two methods in terms of definition and measurement of shares of price discovery. The IS method focuses on the contribution of each market to the innovation of new arrival of information. While the PT method focuses on the adjustment of price series to fit the long-run relationship between two cointegrated series. Therefore, in terms of "variance" measures, namely the IS, the Mainland China market contributes significantly more to price discovery. By comparison, this advantage of Mainland China market is relatively weaker in terms of price adjustment, because results from the PT method show that the Hong Kong market also leads Mainland China for quite a few A-H pairs in certain cointegrated years.

What caused the disparity of shares of price discovery between cross-listing A- and H- shares? We tackle this question in Section 5, by combining our empirical results with market microstructure features and try to find out the reason why the A-shares market dominates most of the price discovery.

## 5. The explanation for IS

Since IS puts more emphasis on determining the contribution to innovation when new information arrives, we try to explain the cross-section and time-series differences of IS for our sample A- and H- shares in their cointegrated years.

Our hypothesis is for Mainland China to have a higher IS because the domestic investors in Mainland China have an information advantage and geographic proximity compared with foreign investors in the Hong Kong stock market. We also expect the various policy changes relating to China's securities regulation to have an influence on the domestic investors' position of informational advantage, hence changing the relative IS of Mainland China.

The literature commencing with French and Poterba (1991) that tracks the information disadvantage of foreign investors, terms this issue the home bias puzzle. It shows that due to an information disadvantage, foreign investors tend to underweight their international component of portfolios. Cooper and Kaplanis (1994) and Tesar and Werner (1995) investigate possible explanations for home bias in different situations. They indicate an information advantage as a potential explanation. In terms of geography proximity for domestic investors (see, for example, Grubel (1968), Levy and Sarnat (1970) and Solnik (1974), and Coval and Moskowitz (1999, 2001)), a possible explanation is that investors geographically in proximity to the firm's headquarters gain an informational advantage. Feng and Seasholes (2004) find highly correlated purchasing and sale behavior for geographically categorized investors in Mainland China. Furthermore, investors living near a firm's headquarters react in a similar manner to releases of public information, indicating that geography proximity is an important factor for the behavior of informed investors. Hong, Kubik and Stein (2005) find that proximity and word-of-mouth effects are important in the transmission of information.

In our paper, we use panel regression analysis to test the hypothesis that the informational advantage of Mainland China domestic investors, represented by the Adverse Selection component of the bid-ask spread, is reflected in the stronger price discovery measure obtained by Hasbrouck's (1995) IS. The cross-sectional effects of industries, stock exchanges, and the time-series effects of policy changes on the IS are also tested. We firstly discuss the dependent variable and the reasons of choosing the explanatory variables, together with the proposed hypothesis. Then, we present and discuss the panel data regression results and their implications.

### 5.1 Dependent variable

The dependent variable is the logistic transformation of the IS (Hasbrouck 1995) of the Mainland China stock market, denoted by *IS_MC*, which is the relative ratio (compared with Hong Kong stock market) of its contribution to price discovery for A-H pair cross-listing stocks. We take the logistic transformation because IS lies between 0 and 1 and it is reasonable to match its range with other explanatory variables.

Since the IS is already a relative ratio, *IS_MC* is not an absolute value, but Mainland China market's relative contribution to informational price discovery compared with the Hong Kong market. For example, for JWT the *IS_MC* in the year 2005 is 0.657, which implies that on average 65.7% of the price discovery happening in the A- and H- shares market for JWT in 2005 was contributed by the Mainland China's A-share market. The average *IS_MC* for our entire sample A-H pairs across their cointegrated years is 0.632.

### 5.2 Explanatory variables

In this section, we discuss the reasons to choose the following explanatory variables and the economic hypothesis behind them.

#### 5.2.1 Bid-Ask spread and its components

We expect *IS_MC* to be inversely related to the relative transaction cost of the Mainland China market compared with the Hong Kong market, because lower trading cost makes it easier for information to be realized and reflected in the equilibrium prices. Harris et al. (2002a) has documented that the NYSE's relative contribution to price discovery, compared with regional exchanges, is negatively related to its relative spreads. In their paper about price discovery of US-Canada cross-listing Canadian stocks, Eun and Sabherwal (2003) also provide evidence that the U.S. share of price discovery is inversely related to its ratio of bid-ask spreads. In this paper, we use effective relative spread to represent the transaction cost in each market. The relative transaction cost of Mainland China market compared with Hong Kong is denoted by the explanatory variable, *EffRelSprd*, the ratio of the average effective relative spread for each sample A-share in each sample year to that of the corresponding H-share in the same year. According to evidences from literature and our economic hypothesis, we expect *EffRelSprd* to be inversely related with *IS_MC*.

The dynamics of development and one of the main contributions of market microstructure research have been the components of bid-ask spread and its decomposition. As was mentioned in the literature review section, Lin et al. (1995), Huang and Stoll (1997), Madhavan et al. (1997) and Kraus and Stoll (1972) try to describe the influence of information on security prices from a microstructure perspective. Therefore, we combine the econometric methods of measuring price discovery with the decomposition of bid-ask spreads, which can be decomposed in to an order processing component, an inventory component and an adverse selection component. We expect that there will be a positive relationship between the IS of one market and its corresponding adverse selection component of the bid-ask spread.

Therefore, we use the Lin et al. (1995) method to decompose the effective relative spread of each sample A-share and H-share stock in each sample year to obtain the Adverse Selection component and the Order Processing components. The explanatory variable *AdvSelComp* is the ratio of the Adverse Selection component in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs. *OrdProcComp* is the ratio of the Order Processing component in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs. Since *AdvSelComp* represents Mainland China's information advantage relative to the Hong Kong market, we expect that there is a significant positive relationship between IS_MC and *AdvSelComp*.

### 5.2.2 Control variables

Since our sample of the dependent variable varies in both cross-section and time-series dimensions, we use dummy variables to control for these effects.

Firstly, we expect that there might be a difference for *IS_MC* across industries. So the A-H pair companies in our sample are divided into eight categories of industries, according to the Shanghai and Shenzhen Stock Exchanges rules, respectively as *Industrials*, *Consumer*, *Materials*, *Energy, Telecommunication, Health, Utilities and Finance.* Therefore, we introduce seven categories of industry dummy variables, as *Industrials*, *Consumer, Materials, Energy, Telecommunication, Health* and *Utilities.*

Secondly, we want to test the hypothesis as to whether stock exchanges have an effect on the price discovery for cross-listing A- and H- shares, because the A-shares in Mainland China markets are listed in either Shanghai Stock Exchange (SSE) or Shenzhen Stock Exchange (SZSE). According to the stock exchange that the A-shares of these cross-listing stocks are listed, respectively the Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE), we introduce a dummy variable *SZSE* for cross-listing stocks whose A-shares are listed in SZSE.

Lastly, for the time-series dimension, we consider two categories of dummy variables. The first category tests for the policy influence of QFII and QDII and we expect that the *IS_MC* to be changed after the carrying out of these two policies. The reason for this hypothesis is that these policies allow financial institutions in Mainland China and abroad to invest in both A- and H- shares markets under certain regulatory restrictions. There has been sufficient deregulation for capital to flow between these two markets to some extent, though not enough to flow freely. The second category of time-series dummy variable is simply the year dummies. Ee expect to see some difference between the relationship of *IS_MC* and year dummies after the year 2006, when the QDII allow domestic banks in Mainland China to invest in the Hong Kong stock market.

### 5.3 Regression results and implications

In this section, we discuss the regression results based on previous analysis and explain the implications for our hypothesis. Most of the regression results are in consistent with our hypothesis expectations. First of all, we present the regression results concerning explanatory variables of most importance, the *EffRelSprd* and its decomposition parts *AdvSelComp* and *OrdProcComp*.

(Insert Table 6 about here)

From Table 6, it is clearly possible to observe a negative relationship between *IS_MC* and *EffRelSprd*, at a 10% significance level. After we decompose the effective relative spread and calculate the ratio for Mainland China market respectively for its two components, we find a significant positive relationship between *IS_MC* and *AdvSelComp*, but no relationship between *IS_MC* and *OrdProcComp*. If we run regression of *IS_MC* only on *AdvSelComp*, the positive relationship becomes even stronger, at a 1% significance level. This is consistent with our expectation that the relative IS of Mainland China compared with Hong Kong is inversely related to the transaction cost represented by its relative ratio of effective relative spread. It is also significantly and positively related to its relative ratio of Adverse Selection component of the effective relative spread. Since the Adverse Selection component represents the information advantage of informed traders, the result implies that the reason for Mainland China having a higher IS than Hong Kong is significantly due to the information advantage of domestic investors in Mainland China A-shares market.

Other than previously summarized literature on the information advantage related to home bias, several recent articles which particularly investigate this issue for cross-listing stocks and international portfolios include Kalev et al. (2008), Bae et al. (2008), Barron and Ni (2008) and Karlsson and Norden (2007). They all examine the information asymmetry between local and foreign investors and provide some evidences that local investors have trading advantages for cross-listed stocks. Hong et al. (2008) test the local bias using data on U.S. states and it is consistent with a home bias and geographic proximity effect for internationally cross-listing stocks. Not only on the investment side is there a “home bias”, Sarkissian and Schill (2004) imply that the same proximity constraints for investment portfolios also exert a profound influence on financing decisions. Therefore, it is reasonable to draw the conclusion that information asymmetry is an important reason for Mainland China having a higher IS.

Next, we present the regression results after controlling for cross-section and time-series effects to test the robustness of our results. Table 7 demonstrates the relations between the dependent variable and explanatory variables after controlling for, respectively, industry and policy dummies, stock exchange and policy dummies, and stock exchange and year dummies. To control and test for both cross-sectional and time-series effects, we add different categories of dummy variables. The A-H pair listing companies in our sample can be divided into eight categories of industries, according to the Shanghai and Shenzhen Stock Exchanges, respectively as Industrials, Consumer, Materials, Energy, Telecommunication, Health, Utilities and Finance. Therefore, we introduce seven categories of industry dummy variables, as *Industrials, Consumer, Materials, Energy, Telecommunication, Health and Utilities*.

For time series dummy variables, we divided the 2001-2007 sample period to three sub-periods according to the QFII and QDII policies by the Chinese government. The QFII and QDII policies are carried out respectively during the year 2002 and the year 2006. Therefore, we introduce two time dummy variables, respectively *PostQFPreQD* (implying time period after the QFII policy in 2002 and before the QDII policy in 2006) and *PostQD* (implying time period after the QDII policy). According to the stock exchange that the A-shares of these cross-listing stocks are listed, respectively the Shanghai Stock Exchange (SSE) and Shenzhen Stock Exchange (SZSE), we introduce a dummy variable *SZE* for cross-listing stocks whose A-shares are listed in SZSE. Lastly, we divide the 2001-2007 sample period into seven sub-periods according to the year itself. Hence, we introduce six time dummy variables, *year1, year2, year3, year4, year5 and year6* respectively for each year from 2001 to 2006.

(Insert Table 7 about here)

From Table 7, the main conclusions with respect to the more fundamental dummy variables, both in cross-section and time-series dimensions, can be drawn. The positive relationship between *IS_MC* and *AdvSelComp* is still significant in all the three regressions. The first regression shows that, for cross-listing companies in certain industries, such as *Industrials*, *Consumer*, *Materials*, *Telecommunication* and *Health*, the IS for Mainland China A-shares market is higher than that of Hong Kong H-shares market. In the second regression, we test for the hypothesis that different stock exchanges will have a different effect on the price discovery. The result strongly support the hypothesis that cross-listing companies whose A-shares are listed in the Shenzhen Stock Exchange (SZSE), have a significantly higher IS than the Hong Kong market, compared with companies listed in Shanghai Stock Exchange (SSE). It means that, for A-shares listed in SZSE, the domestic investors in Mainland China have a greater information advantage than A-shares listed in SSE, compared with H-shares investors in Hong Kong market.

From the third regression, we find that after the year 2006, the IS for Mainland China decreased significantly at the 5% level. In both regression 1 and 2, we do not see any strong relationship between QFII (QDII) policy dummies and IS_MC, so it is reasonable to believe that only QDII (carried out in 2006) has a strong effect on the price discovery for cross-listing A- and H- shares.

## 6. Conclusions

In this paper, we examine the price discovery process for cross-listing A- and H-shares, the shares of a same company simultaneously listed in the Mainland China and the Hong Kong stock market. Our study includes a sample of 30 pairs of A-H shares through 7 years from 2001 to 2007, with 26 A-shares listed in SSE and 4 A-shares listed in SZSE in Mainland China stock market. The sample data include intraday data of 2 hours of overlapping trading time on every trading day for both markets.

Based on a cointegration model, we use the Hasbrouck (1995) IS method and the Gonzalo and Granger (1995) PT method to discuss which market contributes more to the price discovery process for cross-listing A- and H- shares. Our cointegration test year by year, from 2001 to 2007, shows that the prices for Mainland China A-shares market and Hong Kong H-shares market are becoming more and more cointegrated through these seven years. This provides one of the first systematic tests showing the development of Mainland China as an emerging market.

Our empirical results for the IS method provide strong evidence that the Mainland China market plays a predominant role in price discovery for cross listing A- and H-shares, because 97.4% sample pairs show that the A-shares market has a higher IS than the H-shares market. Results of the PT method does not show a formidable dominance for Mainland China market in the price discovery process, but still 72.7% sample pairs indicate that Mainland China market's share of price discovery is higher than that of the Hong Kong market. Consequently, it is reasonable to draw the conclusion that Mainland China market dominates the price discovery for most cross-listing A- and H- shares during their cointegrated years. The results of IS and PT methods do not contradict each other, though they are not exactly the same. It is because that they have different definition and measurements for price discovery and emphasize different dimension of price discovery, with IS focusing on contribution to innovation and PT focusing on price adjustment.

To explore the reason why A-shares market dominates most of the price discovery for cross-listing A- and H- shares, we ran regressions of IS for Mainland China on different explanatory and control variables, which have an economic meaning and are of great interest to policy makers. Our contribution is to combine the econometric methods such as IS with the decomposition of bid-ask spreads from the market microstructure literature.

The regression results are consistent with our expectation and imply a negative relationship between the relative IS of Mainland China and its relative transaction cost represented by its effective relative spread. Most importantly, there is a significantly positive relationship between the relative IS of Mainland China and its relative Adverse Selection component of the effective relative spread. Because the Adverse Selection component represents the information advantage of informed traders, the result implies that the reason for Mainland China having a higher IS than Hong Kong is significantly due to the information advantage and geography proximity of domestic investors in Mainland China A-shares market.

The regression results for controlling dummy variable both in cross-section and time-series dimensions show that, for cross-listing companies whose A-shares are listed in the Shenzhen Stock Exchange (SZSE), they have a significantly higher IS than the Hong Kong market, compared with companies listed in Shanghai Stock Exchange (SSE). We are also interested in the result that after the year 2006, the relative level of IS for Mainland China significantly drops compared with previous years, which is an indication for QDII policy having a strong effect on the investment behavior in the A- and H- shares market.

We are also concerned with further integrating the use of econometric methods in price discovery with key contributions of the market microstructure framework. The influence of the drastic up-and-down change in A-shares market in recent two years is also of great interest. These both represent good directions for future research.

# References

Bae K.H., Stulz R.M., Tan H., 2008, Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts, *Journal of Financial Economics* 88, 581–606.

Baillie R.T., Booth G.G., Tse Y., and Zabotina T., 2002, Price discovery and common factor models, *Journal of Financial Markets* 5, 309–321.

Barron J.M., Ni J., 2008, Endogenous asymmetric information and international equity home bias: The effects of portfolio size and information costs, *Journal of International Money and Finance* 27, 617-635.

Biais B., Martinez I., 2004, Price Discovery across the Rhine, *Review of Finance* 8, 49–74.

Booth G.G., So R.W., Tse Y., 1999, Price discovery in the German equity index derivatives market, *The Journal of Futures Markets*19(6), 619-643.

Booth G.G., Lin J.C., Martikainen T., Tse Y., 2002, Trading and Pricing in Upstairs and Downstairs Stock Markets, *The Review of Financial Studies* 15(4), 1111-1135.

Chakravarty S., Gulen H., and Mayhew S., 2004, Informed Trading in Stock and Option Markets, *The Journal of Finance* 59(3), 1235-1257.

Chan K., Chung P., and Fong W.M., The Informational Role of Stock and Option Volume, 2002, *Review of Financial Studies* 15, 1049-1075.

Chen Z., Swan P.L., 2007, "Liquidity Asset Pricing Model in a Segmented Equity Market," Chapter 23 in Lhabitant, F., Gregoriou, G.N. (Eds), *Stock Market Liquidity: Implications for Market Microstructure and Asset Pricing*, 417-444, Wiley Finance.

Cooper I., Kaplanis E., 1994. Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium. *Review of Financial Studies* 7, 145-160.

Coval J., Moskowitz T., 1999. Home Bias at Home: Local Equity Preference in Domestic Portfolios. *Journal of Finance* 54, 6 2045-2073.

Coval J., Moskowitz,T., 2001. The Geography of Investment: Informed Trading and Asset Prices. *Journal of Political Economy* 109, 4 811-841.

de Jong F., 2002, Measures of contributions to price discovery: a comparison, *Journal of Financial Markets* 5, 323–327.

Domowitz I., Glen J., Madhavan A., 1998, International cross-listing and order flow migration: evidence from an emerging market, *Journal of finance* 53(6), 2001-2027.

Eun C.S., Sabherwal S., 2003, Cross-border listings and price discovery: evidence from U.S.-listed Canadian Stocks, *Journal of Finance* 58(2), 549-575.

Feng L. and Seasholes M.S., 2004, Correlated Trading and Location, *The Journal of Finance* 59( 5), 2117 -2144.

French K., Poterba J., 1991. Investor Diversification and International Equity Markets. *American Economic Review* 81(2), 222-226.

Gonzalo J., and Granger C., 1995, Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business & Economic Statistics* 13(1), 27-35.

Grammig J., Melvin M., Schlag C., 2005, Internationally cross-listed stock prices during overlapping trading hours: price discovery and exchange rate effects, *Journal of Empirical Finance* 12, 139– 164.

Grubel H., 1968. Internationally Diversified Portfolios: Welfare Gains and Capital Flows. *American Economic Review* 58, 1399-1314.

Halling M., Pagano M, Randl O., Zechner J., 2007, Where is the market? evidence from cross-listings in the United States, *Review of Financial Studies* 21, 725-761.

Harris F.H. deB, McInish T.H., Shoesmith G.L., and Wood R.A., 1995, Cointegration, Error Correction, and Price Discovery on Informationally Linked Security Markets, *Journal of Financial and Quantitative Analysis* 30(4), 563 -579.

Harris F.H. deB, McInish T.H., Wood R.A., 2002a. Security price adjustment across exchanges: an investigation of common factor components for Dow stocks. *Journal of Financial Markets* 5, 277–308.

Harris F.H. deB, McInish T.H., Wood R.A., 2002b, Common factor components versus information shares: a reply, *Journal of Financial Markets* 5, 341–348.

Hasbrouck J., Measuring the Information Content of Stock Trades, 1991, *The Journal of Finance* 46(1), 179-207.

Hasbrouck J., The Summary Informativeness of Stock Trades: An Econometric Analysis, 1991, *The Review of Financial Studies* 4(3), 571-595.

Hasbrouck J., 1995, One security, many markets: Determining the contributions to price discovery, *Journal of Finance* 50, 1175-1199.

Hasbrouck J., 2002, Stalking the “efficient price” in market microstructure specifications: an overview, *Journal of Financial Markets* 5, 329–339.

Hasbrouck J., 2003, Intraday Price Formation in U.S. Equity Index Markets, *The Journal of Finance* 58(6), 2375-2399.

Hong H., Kubik J. D., Stein, J. C., 2005, Thy neighbor’s portfolio: Word-of-mouth effects in the holdings and trades of money managers, *Journal of Finance* 60, 2801-2842.

Hong H., Kubik J.D., Stein J.C., 2008, The only game in town: Stock-price consequences of local bias. *Journal of Financial Economics*, forthcoming, doi:10.1016/j.jfineco.2007.11.006.

Huang R.D., Stoll H.R., 1997, The Components of the Bid-Ask Spread: A General Approach, *The Review of Financial Studies* 10(4), 995-1034.

Hupperets E.C.J., Menkveld A.J., 2002, Intraday analysis of market integration: Dutch blue chips traded in Amsterdam and New York, *Journal of Financial Markets* 5, 57-82.

Johansen S., 1991, Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica* 59(6), 1551-1580.

Kalev P.S., Nguyen A.H., Oh N.Y., 2008, Foreign versus local investors: Who knows more? Who makes more? *Journal of Banking and Finance* 32, 2376–2389.

Karlsson A., Norden L., 2007, Home sweet home: Home bias and international diversification among individual investors, *Journal of Banking and Finance* 31, 317–333.

Kraus A., Stoll H.R., 1972, Price Impacts of Block Trading on the New York Stock Exchange, *The Journal of Finance* 27(3), 569-588.

Lehmann B.N., 2002, Some desiderata for the measurement of price discovery across markets, *Journal of Financial Markets* 5, 259–276.

Levy H., Sarnat M., 1970. International Diversification of Investment Portfolios. *American Economic Review* 50, 668-675.

Lin J.C., Sanger G.C., and Booth G.G., 1995, Trade Size and Components of the Bid-Ask Spread, *The Review of Financial Studies* 8(4), 1153-1183.

Madhavan A., Richardson M., Roomans M., 1997, Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks, *The Review of Financial Studies* 10(4), 1035-1064.

Menkveld A.J., Koopman S.J., Lucas A., 2007, Modeling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods, *Journal of Business & Economic Statistics* 25(2), 213-225.

Pascual R., Fuster B.P., Climent F., 2006, Cross-listing, price discovery and the informativeness of the trading process, *Journal of Financial Markets* 9, 144–161.

Sarkissian S., Schill M.J., 2004, The Overseas Listing Decision: New Evidence of Proximity Preference, *The Review of Financial Studies* 17(3), 769-809.

Solnik B., 1974. Why not diversify internationally rather than domestically? *Financial Analysts' Journal* 30, 48-54.

Stock J.H., Watson M.W., 1988, Testing for Common Trends, *Journal of the American Statistical Association* 83(404), 1097 -1107.

Tesar L., Werner I., 1995. Home bias and high turnover. *Journal of International Money and Finance* 14, 4 467-492.

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

Werner I.M., Kleidon A.W., 1996, U.K. and U.S. trading of British cross-listed stocks: an intraday analysis of market integration, *Review of Financial Studies* 9(2), 619-664.

**Table 1: Literature summary and comparison for price discovery in cross-listing countries**

| Literature | Cross-listing countries | Methodology | Main results |
|---|---|---|---|
| Werner and Kleidon (1996) | U.K. and U.S. | Intraday pattern | Domination varies |
| Domowitz et al. (1998) | Mexico and U.S. | Liquidity and volatility | Trade-off impact |
| Hupperets et al. (2002) | Holland and U.S. | Hasbrouck (1995) | Domination varies |
| Eun and Sabherwal (2003) | Canada and U.S. | Hasbrouck modified | Domination varies |
| Wang and Jiang (2003) | Chinese mainland and HK | Daily data regression | Domination varies |
| Biais and Martinez (2004) | Germany and France | Information content | Home dominates |
| Grammig et al. (2005) | Germany and U.S. | Hasbrouck modified | Home dominates |
| Pascual et al. (2006) | Spain and U.S. | Hasbrouck modified | Home dominates |
| Halling et al. (2007) | 34 countries and U.S. | Comparison | Domination varies |
| Menkveld et al. (2007) | Holland and U.S. | State space model | Home dominates |

**Table 2: The sample of 30 pairs of A- and H- shares and their basic information**

| No. | Company Abbr. name | Symbol | A-share code[a] | A-share IPO date | H-share code[a] | H-share IPO date | First Listing[b] |
|---|---|---|---|---|---|---|---|
| 1 | NE ELECTRIC | NEE | 000585.sz | 1995-12-13 | 0042.hk | 1995-7-6 | HK |
| 2 | TSINGTAO BREW | TDB | 600600.ss | 1993-8-27 | 0168.hk | 1993-7-15 | HK |
| 3 | JIANGSU EXPRESS | JSE | 600377.ss | 2001-1-16 | 0177.hk | 1997-6-27 | HK |
| 4 | BEIREN PRINT MACH | BRP | 600860.ss | 1994-5-6 | 0187.hk | 1993-8-6 | HK |
| 5 | KUNMING MACHINE | KMM | 600806.ss | 1994-1-3 | 0300.hk | 1993-12-7 | HK |
| 6 | GUANGZHOU SHIP | GZS | 600685.ss | 1993-10-28 | 0317.hk | 1993-8-6 | HK |
| 7 | MAANSHAN IRON | MASI | 600808.ss | 1994-1-6 | 0323.hk | 1993-11-3 | HK |
| 8 | SHANGHAI PECHEM | SHP | 600688.ss | 1993-11-8 | 0338.hk | 1993-7-26 | HK |
| 9 | JINGWEI TEXTILE | JWT | 000666.sz | 1996-12-10 | 0350.hk | 1996-2-2 | HK |
| 10 | JIANGXI COPPER | JXC | 600362.ss | 2002-1-11 | 0358.hk | 1997-6-12 | HK |
| 11 | SINOPEC CORP | SNP | 600028.ss | 2001-8-8 | 0386.hk | 2000-10-18 | HK |
| 12 | SHENZHEN EXPRESS | SZE | 600548.ss | 2001-12-25 | 0548.hk | 1997-3-12 | HK |
| 13 | NANJING PANDA | NJP | 600775.ss | 1996-11-18 | 0553.hk | 1996-5-2 | HK |
| 14 | CHINA EAST AIR | CEA | 600115.ss | 1997-11-5 | 0670.hk | 1997-2-5 | HK |
| 15 | ZTE | ZTE | 000063.sz | 1997-11-18 | 0763.hk | 2004-12-9 | MC |
| 16 | GUANGZHOU PHARMA | GZP | 600332.ss | 2001-2-6 | 0874.hk | 1997-10-30 | HK |
| 17 | HUANENG POWER | HNP | 600011.ss | 2001-12-6 | 0902.hk | 1998-3-4 | HK |
| 18 | ANHUI CONCH | AHC | 600585.ss | 2002-2-7 | 0914.hk | 1997-10-21 | HK |
| 19 | HISENSE KELON ELEC | HKE | 000921.sz | 1999-7-13 | 0921.hk | 1996-7-23 | HK |
| 20 | ANHUI EXPRESS WAY | AHE | 600012.ss | 2003-1-7 | 0995.hk | 1996-11-13 | HK |
| 21 | YIZHENG CHEM | YZC | 600871.ss | 1995-4-11 | 1033.hk | 1994-3-29 | HK |
| 22 | CHINA SOUTH AIR | CSA | 600029.ss | 2003-7-25 | 1055.hk | 1997-7-31 | HK |
| 23 | TIANJIN CAPITAL | TJC | 600874.ss | 1995-6-30 | 1065.hk | 1994-5-17 | HK |
| 24 | HUADIAN POWER | HDP | 600027.ss | 2005-2-3 | 1071.hk | 1999-6-30 | HK |
| 25 | DONGFANG ELEC | DFE | 600875.ss | 1995-10-10 | 1072.hk | 1994-6-6 | HK |
| 26 | LUOYANG GLASS | LYG | 600876.ss | 1995-10-31 | 1108.hk | 1994-6-29 | HK |
| 27 | CHINA SHIP DEV | CSD | 600026.ss | 2002-5-23 | 1138.hk | 1994-11-11 | HK |
| 28 | YANZHOU COAL | YZC | 600188.ss | 1998-7-1 | 1171.hk | 1998-4-1 | HK |
| 29 | CHINA MERCHANT BANK | CMB | 600036.ss | 2002-4-9 | 3968.hk | 2006-9-22 | MC |
| 30 | BANK OF CHINA | BOC | 601988.ss | 2006-7-5 | 3988.hk | 2006-6-1 | HK |

[a] The A-share code symbol .ss indicates listing in SSE, .sz in SZSE and H-share code symbol .hk indicates listing in HKEx.

[b] This column shows the first listing market for each pair, HK stands for Hong Kong and MC stands for Mainland China.

## Table 3: Estimates and tests of Cointegration over the sample years

For each cross-listing company, we use the 1- minute-interval, mid-quote time-series intraday prices for its A-share and H-share in each year to estimate the Johansen cointegration test statistics. We present the estimation results of maximum Eigen value test of $H_0$: r = 0 against $H_1$: r > 0. The critical value to reject the null hypothesis at the 5% significance level is 15.34. An asterisk indicates rejection of the null hypothesis $H_0$: r = 0. For years through 2001 to 2007, we estimates the years when data is available for sample stocks and calculate the frequency and percentage of significantly cointegrated A-H share pairs for each sample year. We can observe a gradual and obvious increase in the percentage of cointegrated A-H share pairs from the year 2001 to 2007.

| No. | Symbol | Year | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| 1 | NEE | 6.8751 | 16.5155* | 11.2973 | 7.6003 | 10.9231 | 19.3589* | 20.8489* |
| 2 | TTB | | | 13.8998 | 16.8367* | 30.8019* | 4.4996 | 29.0363* |
| 3 | JSE | | | 6.1059 | 16.1464* | 10.0497 | 12.2298 | 19.3741* |
| 4 | BRP | | | 11.391 | 14.7918 | 11.563 | 32.3354* | 30.6553* |
| 5 | KMM | | | 9.2814 | 22.967* | 54.8307* | 15.1142 | 25.2627* |
| 6 | GZS | | | 13.1562 | 11.2154 | 10.7432 | 15.9514* | 21.7893* |
| 7 | MASI | | | 5.244 | 6.7216 | 18.4269* | 17.0302* | 18.0557* |
| 8 | SHP | | | 11.5977 | 9.8632 | 25.1459* | 18.4985* | 21.0191* |
| 9 | JWT | 8.7429 | 5.1591 | 18.1254* | 10.6313 | 24.0619* | 22.36* | 15.3651* |
| 10 | JXC | | | 10.4515 | 6.7958 | 28.6784* | 24.8979* | 21.7916* |
| 11 | SNP | | | 8.1376 | 27.6928* | 16.4912* | 11.3748 | 12.0155 |
| 12 | SZE | | | 9.4036 | 7.0542 | 8.925 | 50.2874* | 19.7572* |
| 13 | NJP | | | 6.0872 | 26.3034* | 7.1919 | 23.9139* | 31.8525* |
| 14 | CEA | | | 16.1654* | 16.2152* | 24.7038* | 8.2298 | 14.2711 |
| 15 | ZTE | 18.4892* | 10.6751 | 39.7597* | 9.3113 | 16.996* | 17.0964* | 21.2273* |
| 16 | GZP | | | 15.5017* | 5.8721 | 15.8875* | 22.1671* | 35.886* |
| 17 | HNP | | | 4.9749 | 15.7588* | 21.2087* | 9.0233 | 17.4064* |
| 18 | AHC | | | 16.8808* | 10.4783 | 28.7881* | 40.6569* | 23.3268* |
| 19 | HKE | 10.2663 | 15.4989* | 19.7737* | 12.9472 | 8.247 | | |
| 20 | AHE | | | 5.7613 | 9.7384 | 7.5483 | 17.0149* | 18.6485* |
| 21 | YZC | | | 6.0706 | 15.332 | 13.4722 | 17.9941* | 30.0842* |
| 22 | CSA | | | 14.4615 | 20.1727* | 32.6493* | 9.024 | 19.4514* |
| 23 | TJC | | | 13.4444 | 8.1395 | 13.7186 | 12.3445 | 18.647* |
| 24 | HDP | | | | | 22.1283* | 6.1637 | 17.5796* |
| 25 | DFE | | | 9.4772 | 6.1873 | 15.8093* | 16.6046* | 19.3863* |
| 26 | LYG | | | 8.2787 | 7.399 | 10.5287 | 28.9011* | |
| 27 | CSD | | | 4.9379 | 11.0197 | 21.6207* | 19.6939* | 24.0951* |
| 28 | YZC | | | 5.8318 | 8.7279 | 8.5508 | 16.3193* | 15.7735* |
| 29 | CMB | | | | | | 7.0942 | 14.4349 |
| 30 | BOC | | | | | | 50.8064* | 16.7506* |
| No. of significant A-H cointegrated pairs | | 1 | 2 | 6 | 8 | 16 | 19 | 25 |
| No. of total A-H pairs | | 4 | 4 | 27 | 27 | 28 | 29 | 28 |
| Frequency of significant A-H cointegrated pairs | | 1/4 | 2/4 | 6/27 | 8/27 | 16/28 | 19/29 | 25/28 |
| Percentage of significant A-H cointegrated pairs | | 25.00% | 50.00% | 22.22% | 29.63% | 57.14% | 65.52% | 89.29% |

**Table 4: Results of IS for cointegrated A-H pairs**

The mid-point of upper and lower bounds of ISs of Mainland China and Hong Kong stock markets, MC stands for Mainland China and HK for Hong Kong. For each A-H pair in each cointegrated year, the ISs of MC and HK sum up to 1, according to our method. Bolded ISs implies Hong Kong's IS higher than 0.5, which means higher than that of Mainland China. We also provide the maximum and minimum values of ISs for each market in each year.

| No. | Symbol | Mid-point ISs (IS) for each A-H pair in each cointegrated year | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2001 | | 2002 | | 2003 | | 2004 | | 2005 | | 2006 | | 2007 | |
| | | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK |
| 1 | NEE | | | 0.691 | 0.309 | | | | | | | 0.681 | 0.320 | 0.658 | 0.342 |
| 2 | TTB | | | | | | | 0.589 | 0.411 | 0.531 | 0.469 | | | 0.628 | 0.372 |
| 3 | JSE | | | | | | | 0.578 | 0.423 | | | | | 0.613 | 0.387 |
| 4 | BRP | | | | | | | | | | | 0.668 | 0.332 | 0.710 | 0.290 |
| 5 | KMM | | | | | | | 0.725 | 0.276 | 0.639 | 0.362 | | | 0.634 | 0.366 |
| 6 | GZS | | | | | | | | | | | 0.587 | 0.413 | 0.568 | 0.432 |
| 7 | MASI | | | | | | | | | 0.679 | 0.322 | 0.546 | 0.455 | 0.598 | 0.402 |
| 8 | SHP | | | | | | | | | 0.620 | 0.381 | 0.553 | 0.448 | 0.633 | 0.368 |
| 9 | JWT | | | | | 0.776 | 0.225 | | | 0.657 | 0.344 | 0.661 | 0.339 | 0.678 | 0.322 |
| 10 | JXC | | | | | | | | | 0.604 | 0.396 | 0.564 | 0.436 | 0.950 | 0.050 |
| 11 | SNP | | | | | | | 0.605 | 0.396 | 0.561 | 0.440 | | | | |
| 12 | SZE | | | | | | | | | | | 0.512 | 0.489 | 0.571 | 0.429 |
| 13 | NJP | | | | | | | 0.541 | 0.460 | | | 0.657 | 0.344 | 0.726 | 0.274 |
| 14 | CEA | | | | | 0.652 | 0.348 | 0.601 | 0.400 | 0.638 | 0.362 | | | | |
| 15 | ZTE | 0.681 | 0.319 | | | 0.650 | 0.350 | | | 0.650 | 0.351 | 0.550 | 0.450 | 0.611 | 0.390 |
| 16 | GZP | | | | | 0.568 | 0.432 | | | 0.639 | 0.361 | 0.598 | 0.403 | 0.678 | 0.322 |
| 17 | HNP | | | | | | | 0.659 | 0.342 | 0.681 | 0.320 | | | 0.575 | 0.426 |
| 18 | AHC | | | | | 0.521 | 0.479 | | | 0.576 | 0.425 | 0.585 | 0.416 | 0.573 | 0.428 |
| 19 | HKE | | | 0.637 | 0.364 | 0.730 | 0.271 | | | | | | | | |
| 20 | AHE | | | | | | | | | | | 0.562 | 0.438 | 0.583 | 0.417 |
| 21 | YZC | | | | | | | | | | | 0.508 | 0.493 | 0.524 | 0.477 |
| 22 | CSA | | | | | | | 0.631 | 0.370 | 0.607 | 0.393 | | | 0.623 | 0.377 |
| 23 | TJC | | | | | | | | | | | | | 0.605 | 0.396 |
| 24 | HDP | | | | | 0.566 | 0.434 | | | 0.636 | 0.365 | | | 0.532 | 0.468 |
| 25 | DFE | | | | | | | | | 0.537 | 0.463 | 0.558 | 0.443 | 0.658 | 0.342 |
| 26 | LYG | | | | | | | | | | | 0.619 | 0.381 | | |
| 27 | CSD | | | | | | | | | 0.608 | 0.392 | 0.509 | 0.492 | 0.581 | 0.420 |
| 28 | YZC | | | | | | | | | | | 0.548 | 0.452 | 0.563 | 0.438 |
| 29 | CMB | | | | | | | | | | | | | | |
| 30 | BOC | | | | | | | | | | | 0.480 | **0.520** | 0.496 | **0.505** |
| Max for each year | | 0.681 | 0.319 | 0.691 | 0.364 | 0.776 | 0.479 | 0.725 | 0.460 | 0.681 | 0.469 | 0.681 | 0.520 | 0.950 | 0.505 |
| Min for each year | | 0.681 | 0.319 | 0.637 | 0.309 | 0.521 | 0.225 | 0.541 | 0.276 | 0.531 | 0.320 | 0.480 | 0.320 | 0.496 | 0.050 |

**Table 5: Results of PT method for cointegrated A-H pairs**

The following table presents the shares of price discovery of Mainland China and Hong Kong using PT method, for each pair of A-H shares in each cointegrated year. The shares of price discovery of MC and HK sum up to 1, according to our method. Bolded shares of price discovery implies Hong Kong's share higher than 0.5, which means higher than that of Mainland China. We also provide the maximum and minimum values of shares of price discovery under PT method for each market in each year.

| No. | Symbol | Shares of price discovery using PT method | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 2001 | | 2002 | | 2003 | | 2004 | | 2005 | | 2006 | | 2007 | |
| | | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK | MC | HK |
| 1 | NEE | | | 0.748 | 0.252 | | | | | | | 0.781 | 0.219 | 0.787 | 0.213 |
| 2 | TTB | | | | | | | 0.684 | 0.316 | 0.480 | **0.520** | | | 0.537 | 0.463 |
| 3 | JSE | | | | | | | 0.616 | 0.384 | | | | | 0.744 | 0.256 |
| 4 | BRP | | | | | | | | | | | 0.849 | 0.151 | 0.724 | 0.276 |
| 5 | KMM | | | | | | | 0.851 | 0.149 | 0.806 | 0.194 | | | 0.942 | 0.058 |
| 6 | GZS | | | | | | | | | | | 0.764 | 0.236 | 0.720 | 0.280 |
| 7 | MASI | | | | | | | | | 0.443 | **0.557** | 0.161 | **0.839** | 0.601 | 0.399 |
| 8 | SHP | | | | | | | | | 0.470 | **0.530** | 0.619 | 0.381 | 0.818 | 0.182 |
| 9 | JWT | | | | | 0.948 | 0.052 | | | 0.326 | **0.674** | 0.378 | **0.622** | 0.753 | 0.247 |
| 10 | JXC | | | | | | | | | 0.248 | **0.752** | 0.324 | 0.676 | 0.724 | 0.276 |
| 11 | SNP | | | | | | | 0.855 | 0.145 | 0.403 | **0.597** | | | | |
| 12 | SZE | | | | | | | | | | | 0.371 | **0.629** | 0.762 | 0.238 |
| 13 | NJP | | | | | | | 0.853 | 0.147 | | | 0.674 | 0.326 | 0.618 | 0.382 |
| 14 | CEA | | | | | 0.708 | 0.292 | 0.997 | 0.003 | 0.776 | 0.224 | | | | |
| 15 | ZTE | 0.884 | 0.116 | | | 0.988 | 0.012 | | | 0.422 | **0.578** | 0.785 | 0.215 | 0.742 | 0.258 |
| 16 | GZP | | | | | 0.909 | 0.091 | | | 0.909 | 0.091 | 0.890 | 0.110 | 0.941 | 0.059 |
| 17 | HNP | | | | | | | 0.018 | **0.982** | 0.262 | **0.738** | | | 0.521 | 0.479 |
| 18 | AHC | | | | | 0.714 | 0.286 | | | 0.363 | **0.637** | 0.381 | **0.619** | 0.940 | 0.060 |
| 19 | HKE | | | 0.703 | 0.297 | 0.995 | 0.005 | | | | | | | | |
| 20 | AHE | | | | | | | | | | | 0.587 | 0.413 | 0.364 | **0.636** |
| 21 | YZC | | | | | | | | | | | 0.507 | 0.493 | 0.549 | 0.451 |
| 22 | CSA | | | | | | | 0.602 | 0.398 | 0.381 | **0.619** | | | 0.720 | 0.280 |
| 23 | TJC | | | | | | | | | | | | | 0.817 | 0.183 |
| 24 | HDP | | | | | | | | | 0.281 | **0.719** | | | 0.916 | 0.084 |
| 25 | DFE | | | | | | | | | 0.592 | 0.408 | 0.947 | 0.053 | 0.953 | 0.047 |
| 26 | LYG | | | | | | | | | | | 0.706 | 0.294 | | |
| 27 | CSD | | | | | | | | | 0.345 | **0.655** | 0.995 | 0.005 | 0.452 | **0.548** |
| 28 | YZC | | | | | | | | | | | 0.215 | **0.785** | 0.510 | 0.490 |
| 29 | CMB | | | | | | | | | | | | | | |
| 30 | BOC | | | | | | | | | | | 0.434 | **0.566** | 0.551 | 0.449 |
| Max for each year | | 0.884 | 0.116 | 0.748 | 0.297 | 0.995 | 0.292 | 0.997 | 0.982 | 0.909 | 0.752 | 0.995 | 0.839 | 0.953 | 0.636 |
| Min for each year | | 0.884 | 0.116 | 0.703 | 0.252 | 0.708 | 0.005 | 0.018 | 0.003 | 0.248 | 0.091 | 0.161 | 0.005 | 0.364 | 0.047 |

**Table 6: Regression results using components of Bid-Ask spread**

The dependent variable *IS_MC* is the Logistic transformation of the IS (Hasbrouck 1995) of the A-shares in Mainland China stock market, which is the relative ratio (compared with Hong Kong stock market) of its contribution to price discovery for A-H pair cross-listing stocks. *EffRelSprd* is the ratio of the average effective relative spread for each sample A-share in each sample year to that of the corresponding H-share in the same year. We use the Lin et al. (1995) method to decompose the effective relative spread (*EffRelSprd*) of each sample A-share and H-share stock in each sample year and get the Adverse Selection component and the Order Processing component. The explanatory variable *AdvSelComp* is the ratio of the Adverse Selection component in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs in each year. *OrdProcComp* is the ratio of the Order Processing component in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs in each year. There are 76 observations used in this panel data regression, including 29 A-H pair companies across 7 years. The *t-statistics* are in the parentheses below the coefficients. Significance at the 1, 5, 10 percent levels is indicated by ***, ** and * respectively.

| | (1) | (2) | (3) |
|---|---|---|---|
| Intercept | 0.64***<br>(32.65) | 0.60***<br>(62.93) | 0.60***<br>(65.82) |
| *EffRelSprd* | -1.87*<br>(-1.62) | | |
| *AdvSelComp* | | 0.10**<br>(2.11) | 0.11***<br>(2.59) |
| *OrdProcComp* | | 0.04<br>(0.61) | |
| Adjusted R-square (%) | 2.14 | 5.26 | 6.07 |

**Table 7: Regression results after adding different categories of dummy variables**

The dependent variable *IS_MC* is still the Logistic transformation of the IS (Hasbrouck 1995) of the Mainland China stock market, which is the relative ratio (compared with Hong Kong stock market) of its contribution to price discovery for A-H pair cross-listing stocks. The explanatory variable *AdvSelComp* is the ratio of the Adverse Selection component (Lin et al. (1995)) in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs in each year. *OrdProcComp* is the ratio of the Order Processing component (Lin et al. (1995)) in Mainland China market to that of the Hong Kong market for each corresponding A-H pairs in each year. We introduce seven categories of industry dummy variables, as *Industrials*, *Consumer*, *Materials, Energy*, *Telecommunication, Health* and *Utilities.* For time series dummy variables, we introduce two time dummy variables, respectively *PostQFPreQD* (implying time period after the QFII policy in 2002 and before the QDII policy in 2006) and *PostQD* (implying time period after the QDII policy). According to the stock exchange that the A-shares of these cross-listing stocks are listed, we introduce a dummy variable *SZE* for cross-listing stocks whose A-shares are listed in SZSE. Lastly, we divide the 2001-2007 sample period to seven sub-periods according to the year itself. Hence, we introduce 6 time dummy variables, *year1*, *year2*, *year3*, *year4*, *year5 and year6* respectively for the years from 2001 to 2006. There are 76 observations used in these 3 regressions, including 29 A-H pair companies across 7 years. The 3 regressions are run on 3 different categories of dummy variables presented in Table 7. The *t-statistics* are shown in the parentheses below the coefficients. Significance at the 1, 5, 10 percent levels is indicated by ***, ** and * respectively.

| | (1) | (2) | (3) |
|---|---|---|---|
| Intercept | 0.49*** | 0.55*** | 0.62*** |
| | (6.52) | (9.47) | (41.87) |
| *AdvSelComp* | 0.12** | 0.11** | 0.13** |
| | (2.07) | (2.09) | (2.07) |
| *OrdProcComp* | 0.04 | 0.03 | -0.03 |
| | (0.46) | (0.38) | (-0.27) |
| *Industrials* | 0.12** | | |
| | (2.36) | | |
| *Consumer* | 0.15*** | | |
| | (2.72) | | |
| *Materials* | 0.12** | | |
| | (2.25) | | |
| *Energy* | 0.10 | | |
| | (1.66) | | |
| *Telecommunication* | 0.13** | | |
| | (2.34) | | |
| *Health* | 0.13** | | |
| | (2.14) | | |
| *Utilities* | 0.11* | | |
| | (1.84) | | |
| *PostQFPreQD* | -0.01 | 0.06 | |
| | (-0.25) | (1.08) | |
| *PostQD* | 0.02 | 0.03 | |
| | (0.32) | (0.53) | |
| *SZE* | | 0.06*** | 0.06*** |
| | | (2.67) | (2.50) |
| *year1* | | | 0.03 |
| | | | (0.27) |
| *year2* | | | -0.09 |
| | | | (-1.22) |
| *year3* | | | -0.03 |
| | | | (-0.83) |
| *year4* | | | -0.03 |
| | | | (-0.95) |
| *year5* | | | -0.01 |
| | | | (-0.62) |
| *year6* | | | -0.05** |
| | | | (-2.48) |
| Adjusted R-square (%) | 8.09 | 14.10 | 14.16 |

**Figure 1: The trading hours for Mainland China and Hong Kong stock markets**

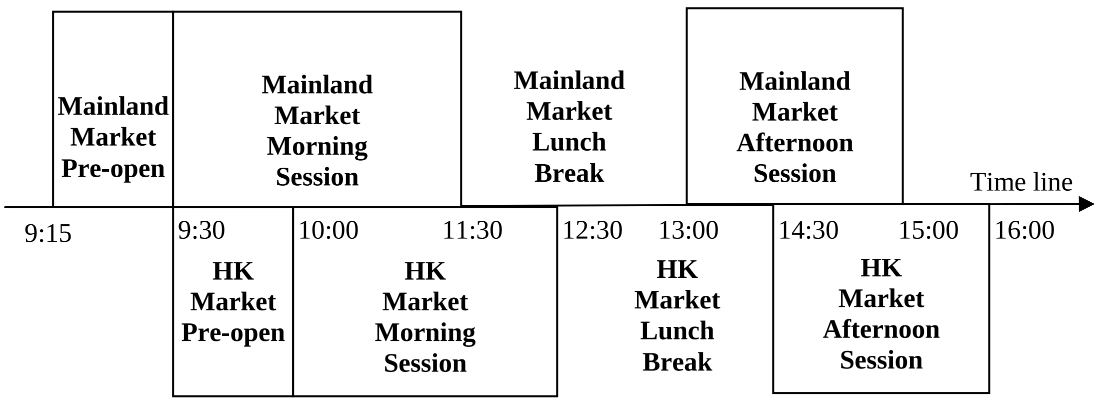
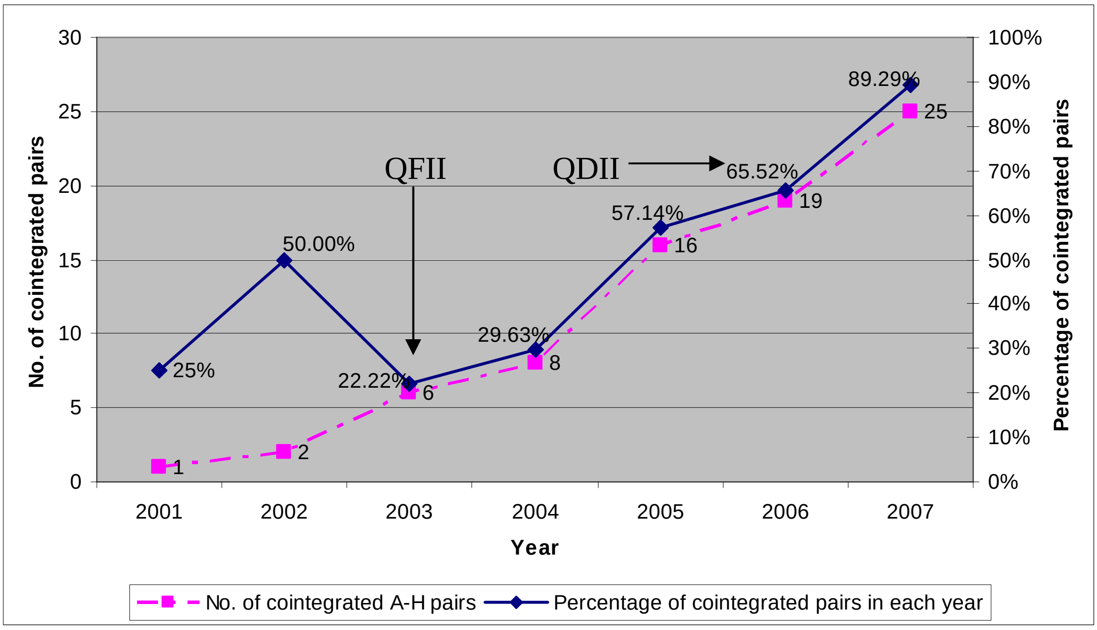


**Figure 2: Evolution trend of cointegrated A- and H-shares pairs in each year**




The dashed line represents the absolute number of cointegrated A-H pairs in each year from 2001 to 2007, while the solid line shows the percentage of cointegrated A-H pairs out of all available samples in each year.

**Figure 3: The scattering of Information Shares for A-H pairs in their cointegrated years**

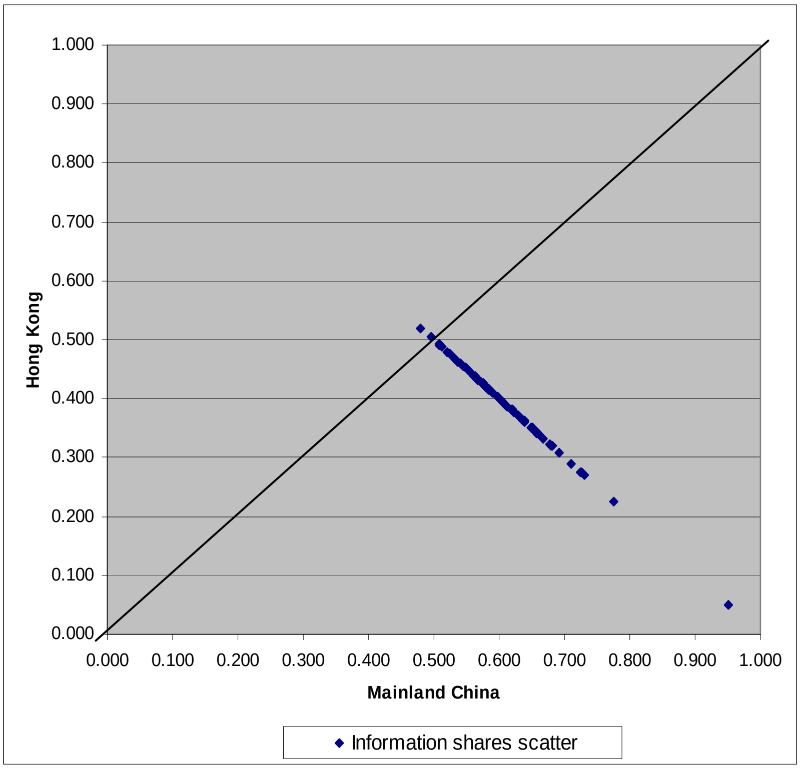


**Figure 4: The scattering of Shares of price discovery for A-H pairs in their cointegrated years using the PT method**