This is a repository copy of *Pairs trading across Mainland China and Hong Kong stock markets*.

Article:

Pairs Trading Across Mainland China and Hong Kong Stock Markets

Hanxiong Zhang[1] and Andrew Urquhart[2]

Abstract

Motivated by the rationale that market inefficiency arises from a combination of less than fully rational demand and limits to arbitrage, this paper investigates the profitability of pairs trading across mainland China and Hong Kong on highly-liquid large and mid-cap stocks from January 1996 to July 2017. We have three main findings. First, we find that pairs trading constrained within each market generates no significant abnormal returns. However, if investors can trade across mainland China and Hong Kong, pairs trading is profitable after adjusting for risk and transaction costs, where the annualized abnormal return is 9% over the full sample. Second, by using a rolling-window regression, we find the profitability of the strategy is time-varying. The bootstrap simulations suggest that the decline in profitability of the strategy since 2012 is due to random chance rather than poor ability of identifying mispriced stocks. However the vast majority of profitable periods reflect the strategy's ability to choose profitable stocks rather than random chance. Third, the profitability of the strategy is somewhat sensitive to market conditions, most notably, the strategy is more profitable during longer-term market turbulence. Overall, our empirical findings are consistent with the Adaptive Market Hypothesis in that the integration of financial markets and market conditions determine the level of market efficiency.

Keywords: Pairs Trading, Bootstrap Simulation; Adaptive Market Hypothesis; Market Efficiency; Cross Border Trading

JEL Classification: G10, G11

---

[1] Lincoln International Business School, University of Lincoln, Brayford Pool, Lincoln, United Kingdom, hzhang@lincoln.ac.uk

[2] Centre of Digital Finance, Southampton Business School, University of Southampton, Southampton, United Kingdom, a.j.urquhart@soton.ac.uk

1. Introduction

Motivated by the theoretical hypothesis and empirical findings that market inefficiency arises from some combination of less than fully rational demand and limits to arbitrage (Baker *et al.*, 2014; Jacobs and Weber, 2015), this paper studies pairs trading from three perspectives. First, we study whether pairs trading across mainland China and Hong Kong stock markets is more profitable than trading in the individual market by using highly-liquid stocks. Second, whether the profitability of pairs trading declines since the higher integration between mainland China and Hong Kong stock markets. Third, what drives this profitability and whether this profitability is time-varying.

Pairs trading is a popular statistical arbitrage strategy used by hedge funds and investment banks since 1987 (Gatev *et al.*, 2006). Essentially the strategy finds two stocks whose prices move together over time and if the pair of prices diverge wide enough, investors buy the declining stock and sell the increasing stock simultaneously to take the advantage of the mispriced stocks while maintain a level of market-neutrality. The rationale behind pairs trading is to profit from mean-reversion forces that eliminate short-term price deviations in favour of long-term historical pricing relationships. There have been a number of papers that have examined the profitability of pairs trading. Gatev *et al.* (2006) find that pairs trading yields statistically significant monthly excess return in the order of 0.9% and a risk-adjusted return of 0.76% before transaction costs in the US stock market by using the CRSP stocks from 1962 to 2002. After considering market impact, proxied by bid-ask spread, the monthly profits drop to 0.19% to 0.38%. To rule out the concern that their findings are data-mining, Gatev *et al.* (2006) conduct an out-of-sample and show that pairs trading remains profitable. Using CRSP data from 1962 to 2009, Do and Faff (2010) find that the profitability of pairs trading peaks in 1970s and 1980s, and declines since 1990s. Do and Faff (2012) find that the strategy performs strongly during periods of market turbulence and after adjusting for commissions, market impact, and short-selling fees, they find the results of Gatev *et al.* (2006) loses its profitability. Nevertheless, pairs trading remains profitable in a fairly small number of refined versions and at much diminished levels. Gatev *et al.* (2006), Do and Faff (2010) and Do and Faff (2012) focus on the highly-liquid US stocks by including stocks in the CRSP daily files that are alive throughout the 18 month period (12 month formation period and 6 month trading period). When a stock in a pair is delisted from CRSP, Gatev *et al.* (2006) closes the position in that pair, using the delisting return, or the last available price. Gatev *et al.* (2006) find that more than 90% of stocks in the pairs trading come from the top five size deciles using CRSP breakpoints. Do and Faff (2012) excludes stocks in the bottom size decile using New York Stock Exchange breakpoints, and those with closing prices in the one-year formation period less than $5. Broussard and Vaihekoski (2012) extend the research to the illiquid Finnish market and find the strategy to be profitable even allowing for a 1-day delay in the trade execution after the trading signal. Jacobs and Weber (2015) find the strategy is persistently profitable in 35 international stock markets and that the strategy is quite successful for pairs that are hard to arbitrage or less visible. Using the constituents of FTSE All-Share Index from 1979 to 2012, Bowen and Hutchinson (2016) find the pairs trading performs well in market turmoil, and argue that the abnormal returns in the UK can be accounted for by a combination of time-varying risk exposures and transaction costs.

A number of different methodologies have been employed to take advantage of pairs trading. For example, Elliott *et al.* (2005) propose a Gaussian-Markov chain model for the spread. Tourin and Yan (2013) propose a dynamic pairs trading strategy using the stochastic control approach. In the spirit of Vidyamurthy (2004)'s cointegration method, Li *et al.* (2014) find the pairs trading yields an average annualized excess return of about 17.6% using 38 firms dual listed on A-shares market in mainland China and H-shares in Hong Kong during the period 2009-2013 while Marshall *et al.* (2013) find the pairs trading is profitable and support the Adaptive Market Hypothesis (AMH) by using two extremely liquid S&P500 ETFs. Xie *et al.* (2016) find the copula method is superior to distance method by investigating 89 US stocks in the utility industry with a sample from 2003 to 2012. Rad *et al.* (2016) investigate the performance of the distance method, cointegration method and copula method on the US stock market from 1962 to 2014 and find that all 3

methods show significant profitability, however, the distance method outperforms the other methods slightly after adjusting for transaction costs.[3]

Profitable pairs trading however, is a challenge to Efficient Market Hypothesis (EMH) which in the weak form, states that prices already reflect all information that can be derived by examining past market trading data such as the history of past prices, trading volume etc. If prices are predictable and profits could be made by using historical data, arbitrage would eliminate these profits in an efficiently operating market. Therefore there should be no predictability in security prices. However successful pairs trading employs past data to predict future prices and therefore violates the EMH. From a theoretical perspective, Grossman and Stiglitz (1980) argue that a perfectly efficient market is impossible if prices correctly reflect all available information, no one would have any motivation to acquire costly information. Campbell *et al.* (1997) suggest market efficiency is not an all-or-nothing absolute condition but a relative notion. As a modification to EMH, Lo (2004) proposes the Adaptive Market Hypothesis (AMH). The AMH extends the EMH view of the market to argue that learning, competition, and evolutionary selection pressures govern the forces that drive prices to their efficient levels. The AMH provides a number of practical implications within finance. Firstly, the risk premium varies over time according to the stock market environment and the demographics of investors in that environment. The second implication is that arbitrage opportunities do exist from time to time in the market. Thus from an evolutionary viewpoint, active liquid financial markets imply that profit opportunities must exist. However as they are exploited, they disappear. But new opportunities are continually being created as certain species/traders die out and rather than move towards a higher degree of efficiency the AMH implies that complex market dynamics such as trends, panics, bubbles and crashes are continually witnessed in natural market ecologies. The third implication is that investment strategies are successful or unsuccessful, depending on the particular market environment. Contrary to the EMH, the AMH implies that investment strategies may decline for a time, and then return to profitability when environmental conditions become more conducive to such strategies. A consequence of this implication is that market efficiency is not an all-or-nothing condition, but is a characteristic that varies continuously over time and across markets. Lo (2005) argues that convergence to equilibrium is neither guaranteed nor likely to occur and that it is incorrect to assume that the market must move towards some ideal state of efficiency. The AMH has gained quite a lot of attention in the recent literature such as Kim *et al.* (2011), Urquhart and Hudson (2013), Urquhart and McGroarty (2014) and Hiremath and Narayan (2016).[4]

Since the original paper by Gatev *et al.* (2006) the literature has expanded but the issue of cross-border pairs trading is not well developed. This paper thereby makes two main contributions to the literature. Firstly, we are the first to study pairs trading by examining highly-liquid pairs trading in the Chinese stock market from 1996 to 2017 across mainland China and Hong Kong while also evaluating the performance of pairs trading over time to determine whether pairs trading evolves over time in a manner similar to the AMH. There are two reasons that make the Chinese stock market attractive for studying whether pairs trading across markets is more profitable. First, the institutional characteristic regarding dual listing on mainland China and Hong Kong stock exchanges provides a unique opportunity to investigate potential arbitrage provided not only by similar stocks, but by stocks that have access to the same cash-flow source (Broussard and Vaihekoski, 2012) which may reduce the risk of non-convergence after a pair opens. Second, the historically poor integration of mainland China and Hong Kong stock markets allows us to study whether the integration of financial markets through raising mutual market access helps to improve market efficiency in both mainland China and Hong Kong markets which, in turn, diminishes the profitability of pairs trading. Stiglitz (2003) and Baele (2005) argue that allowing unfettered flows of speculative capital across markets is extremely risky, while Bae *et al.* (2006) and Sun *et al.* (2009) argue that integration of financial markets tends to improve information environment and enhance market efficiency. Although one other paper to our knowledge (Li *et al.*, 2014) studies pairs trading exclusively in the Chinese stock market, we modify the research design in the following ways in order to add to the current literature: (1), Li *et al.* (2014) use the

---

[3] For a recent review of the literature on the pairs trading, see Krauss (2017).
[4] For a recent review of the literature on the AMH, see Urquhart and McGroarty (2016).

cointegration method to identify the pairs, whereas we use the distance method proposed by Gatev *et al.* (2006). Cointegration is a long-run phenomenon that needs long spans of data (Hakkio and Rush, 1991) and empirical tests on daily data over a year span of time may not reveal a real cointegrated relationship (Gatev *et al.*, 2006; Broussard and Vaihekoski, 2012). (2) Li *et al.* (2014) uses 38 companies dual listed on the A-share market in mainland China and the H-share market in Hong Kong, whereas we use 4 distinct datasets exclusively covering large- and mid-cap stocks. (3) Li *et al.* (2014) use an all-or-nothing static view to study whether pairs trading is profitable, whereas we use a dynamic view to study whether the profitability of pairs trading is time-varying and what factors drive the time-varying profitability. Appreciating the impact of these methodological differences is important to our knowledge of pairs trading. Moreover, each of these differences suggests a significant change in the implementation of trading strategies in practice. Complementing Li *et al.* (2014), our results show over the full sample period pairs trading is unprofitable when investors are constrained to trading in either mainland China or Hong Kong stock market. However once investors can trade across mainland China and Hong Kong, the strategy is profitable where the annualized abnormal return can be up to 9% after transaction costs. Consistent with Gatev *et al.* (2006), we find there is no clear relationship between profitability and number of pairs in a portfolio. By regressing monthly net excess return series against Fama and French (2015) 5-factor plus momentum and short-term reversal factors with 120-month window length rolling forward 1-month each time, we find the profitability of pairs trading is time-varying and disappears after 2012.

Secondly, unlike the existing literature, which examine out-of-sample profitability (Gatev *et al.*, 2006; Do and Faff, 2012; Bowen and Hutchinson, 2016) or adjusted significance levels (White, 2000), we use a modified version of Fama and French (2010) bootstrap simulation to study whether the profitability of pairs trading is due to data-mining (good luck) or real stock picking ability. Our bootstrap simulation eliminates the possibility that luck (either good or bad) is the main reason for our results. The bootstrap simulation does not impose an assumption that returns are collected from any specific parametric distribution, nor does it depends on large-sample asymptotic theory (Kosowski *et al.*, 2006). This makes the simulation well-fitted for the Chinese market with a limited number of monthly returns. Furthermore, we measure the performance distribution of the "best subsample", not by resampling from the distribution of the best subsample *ex-post*, given this ignores the other luck distributions encountered by all other subsamples, these other luck distribution offer highly useful and relevant information (Cuthbertson *et al.*, 2008). We find those poorly performed subsamples such as the declining of profitability in recent years is due to bad luck rather than the poor ability of the strategy and that the majority of outperformances are the result of the strategy's superior arbitrage ability.

We find pairs trading strategy in China is somewhat sensitive to market conditions especially longer-term market conditions. Consistent with Do and Faff (2010) and Bowen and Hutchinson (2016), we also find that pairs trading performs well during longer-term market downturns when investors trade across mainland China and Hong Kong stock markets. Furthermore, we find pairs trading could be a useful return enhancer for those rational risk-averse investors by using the Treynor and Black (1973) portfolio allocation strategy. Overall, our findings are consistent the AMH since we document the time-varying profitability of pairs trading and that the arbitrage opportunity across mainland China and Hong Kong stock markets can be interpreted as "food source on which market participants depend for their survival" (Lo, 2004; Marshall *et al.*, 2013).

The rest of the paper is presented as follows. Section 2 reviews the Chinese stock markets while Section 3 shows the data and empirical methodology. Section 4 discusses the empirical findings. Section 5 concludes the paper.

2. Chinese Stock Market

This section briefly reviews the main characteristics and evolve of the Chinese stock market since it is institutionally different to the stock markets in developed countries and understudied in the literature. The Shanghai Stock Exchange (SHSE) and the Shenzhen Stock Exchange (SZSE), mainland China's two major

stock exchanges, established in December 1990 and April 1991, respectively. The Stock Exchange of Hong Kong (SEHK) was formally established in 1891.[5] By 2016, the SHSE and the SZSE have about 1,200 and 1,800 stocks, respectively; SEHK has about 2,000 listed stocks and half of them are from mainland China. In terms of market capitalization, all 3 stock exchanges ranked top 10 in the world.

Firms incorporated in mainland China can issue different types of common stock depending on where they are listed and which investors are permitted to trade them. The most typical types are A-, B- and H-shares, which are all Chinese Yuan (also named Renminbi, official currency in mainland China) denominated stocks but traded in different currencies, depending on where they are listed. A-shares are stocks, quoted in Chinese Yuan, exclusively available to mainland China citizens and home institutions until 2003; thereafter, also available to non-resident investors under the Qualified Foreign Institutional Investor (QFII), the Renminbi Qualified Foreign Institutional Investor (RQFII)[6], or via the Stock Connect schemes. B-shares, quoted in US Dollars on SHSE and Hong Kong Dollars on SZSE, are exclusively available to foreign investors until 19/02/2001; thereafter, also available to mainland China investors. H-shares, quoted in Hong Kong Dollars, are listed on SEHK. Like other stocks listed on Hong Kong, there are no additional restrictions on who can trade H-shares. Firms could choose to list their A-shares (or B-shares) on ether the SHSE or SZSE, but not on both exchanges. Firms could issue both A-shares and B-shares (or H-shares). Chinese firms that are listed as B- and H-shares are generally subject to stricter disclosure requirements, therefore, they are normally more financially stable than A-shares. Given that B- and H-shares markets are dominated by well-informed foreign institutional investors while A-shares are traditionally dominated by domestic non-professional retail investors, it is widely believed that investors in B- and H-shares will be more rational than their counterparts in A-shares. Unlike the A- and H-shares markets, the B-shares market had never been popular due to the very poor liquidity (Financial Times, 2013). Accordingly, stocks listed on Hong Kong should reflect economic fundamentals better and more integrated with the world financial markets than the A- and B-shares listed on mainland China.

On 17/11/2014, Shanghai and Hong Kong stock exchanges launch the Stock Connect scheme that connects the Shanghai and Hong Kong's stock markets with the aim of raising mutual market access by relaxing restrictions that historically isolate the mainland China and Hong Kong stock markets. The programme extends to encompass the Shenzhen market on 05/12/2016. The schemes would create a single China stock market that ranks as one of the largest in the world by market cap and daily trading turnover. Some people would expect these mutual market access schemes will boost trading volumes and help investors to arbitrage across markets which, in turn, may enhance market efficiency over time (Sun *et al.*, 2009). Given the relative small quotas for the schemes, small number of eligible stocks, differences in trading hours, clearing and settlement systems and holidays[7], the mutual market access is only partial integration by now. Table A2 in Appendices presents a number of key dates in the Chinese stock markets.

3. Data and Empirical Methodology

3.1 Data

We collect daily data over the period 02/01/1996 to 30/07/2017 from four datasets to investigate whether the integration of financial markets matters to the market efficiency. Specially, the Chinese Securities Index

---

[5] See https://www.hkex.com.hk/eng/stat/statrpt/factbook/documents/fb99_01.pdf.

[6] See http://english.sse.com.cn/investors/qfii/what/

[7] Northbound (for Hong Kong investors) and southbound (for mainland China investors) daily trading quotas are set at 13 and 10.5 billion Chinese Yuan, respectively, which account for no more than 5% of daily trading volume in Chinese stock markets by 2016. In the same time zone, opening price for a security in SHSE and SZSE is available at 9:25 while opening price in SEHK is available at 9:20. SHSE and SZSE close at 15:00, while SEHK close at 16:00 on normal trading days. See more detail at http://www.hkex.com.hk/eng/csm/securities_information_book_and_frequently_asked_questions.htm and http://english.sse.com.cn/investors/shhkconnect/mechanism/rules/

300[8] (CSI300) constituents, which consist of the largest 300 stocks (282 stocks in our sample[9]) traded on A-shares market in mainland China (either on Shanghai Stock Exchange or Shenzhen Stock Exchange). The Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents, which consist of 97 large- and 164 mid-cap H-shares and other stocks (261 in total) listed on Stock Exchange of Hong Kong by 25/11/2016. The Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents, which include stocks of firms dual-listed on A-shares market in mainland China and H-shares market in Hong Kong. HSAHP constituents are comprised of 126 stocks for 63 distinct firms by the end of 2016. The last dataset, named ALL constituents, comprised of the first 3 datasets and contain 622 distinct stocks. HSHKI and HSAHP constituents are eligible for southbound trading under the Stock Connect schemes and make them investable reference sets for mainland China investors. These large- and mid-cap stocks are highly-liquid that ensure the applicability of our pairs trading strategy and low transaction costs which, in turn, alleviates the nonsynchronous trading bias.[10] To alleviate survivorship bias, we include all the dead stocks in the dataset up to the date they are delisted from the listed exchange.

Although the offical openensss of A-shares market begins in 2003, across markets trading is implementable to mainland Chinese investors especially institutional investors throughout our sample. This is because a rising number of mainland China brokerage houses such as Guotai Junan[11] operates comprehensive securities trading businesses in Hong Kong since 1995, and there is no restriction on who can trade shares in Hong Kong. Therefore, a mainland China investor can open two accounts to trade stocks across mainland China and Hong Kong simultaneously. Short-selling was officially permitted in A-shares market since 31/03/2010 for a number of large- and mid-cap stocks. Before that, short-selling could take place as an Over-The-Counter (OTC) trade between two traders. It is also possible that within a securities firm, different departments could lend stocks back and forth from other departments (Broussard and Vaihekoski, 2012). Stocks under the Stock Connect schemes are eligible for Margin Trading and Short-Selling. By the end of 2013, more than 700 A-shares are eligible for short-selling.[12]

We use the value-weighted A-shares stock returns to proxy the broad market return, and use 3-month China time deposit rate to proxy the risk-free rate. We use total return indices to proxy the stock prices, which adjust for the dividends, issues, splits and other corporate finance actions. All of this data is sourced from Datastream.

3.2 Implementation of pairs trading

The implementation of pairs trading strategy involves proceeding in two periods, namely, a formation period and a trading period. Following Gatev *et al.* (2006) and Broussard and Vaihekoski (2012), we find pairs of stocks during a 12-month formation period by using the Sum of Squared Deviation based distance method, and trade during a 6-month period immediately following the formation period.

At the beginning of each formation period, the index of stock prices for each stock is normalized to start from 1 as shown in Equation (1).

$$P^*_{i,t;t+1;t+2;\dots;T} = \frac{P_{i,t;t+1;t+2;\dots;t+T_F}}{P_{i,t}} \qquad (1)$$

Where, $P_{i,t}^*$ is the normalized price of stock i on day t. $P_{i,t}$ is the price series of stock i on day t. T_F is the number of trading days in the formation period. Pairs are identified for trading by matching each stock i with a second stock j that has the smallest Sum of Squared Deviation ($SSD_{i,j}$) between the two normalized price series over the 12-month formation period.

$$SSD_{i,j} = \sum_{t=1}^{T_F} \left(P_{i,t}^* - P_{j,t}^* \right)^2 \qquad (2)$$

Top N pairs, ranked by smallest Sum of Squared Deviations ($SSD_{i,j}$), are selected at the end of each formation period and are traded over the next following 6-month trading period.

The trading period starts on the first trading day following the end of the formation period. During the trading period, the pair trade is opened when the normalized stock prices diverge by more than 2 historical standard deviations of the price difference during the formation period. The pair trade is opened by purchasing 1 Chinese Yuan (mainland China currency) in the stock with lower normalized price, and selling 1 Chinese Yuan in the stock with higher normalized price. The long-short trades yield a net position of 0. Thereby, the trade is regarded self-financed. The pair trade is closed when the normalized price series converge, or on the last day in the trading period whether or not price convergence happened. We apply pairs trading under two different trading scenarios: (1) all trades are executed immediately at the end of the day when the signal appears[13], and (2) all trades that have a 1-day delay. The rationale for executing the pairs with a 1-day delay is that some investors, especially non-professional retail investors, often react to the trading signals with a delay. Moreover, it alleviates concerns regarding the bid-ask spread in the market and considers potential difficulties in executing the trade such as nonsynchronous trading. We also require the trading volume above 0 for executing a trade.

To calculate the return for a pair of stocks throughout the trading period, we accumulate weighted daily returns from the long and short positions. The value-weighted daily percentage returns for a pair ($R_{p,t}$) of two stocks is calculated as:

$$R_{p,t} = w_{L,t} r_{L,t} - w_{S,t} r_{S,t} \qquad (3)$$

Where, $r_{L,t}$ and $r_{S,t}$ are the daily percentage returns for the long position and short position and the weights for both stocks ($w_{L,t}$ and $w_{S,t}$) are set to start from 1 after which they change according to the moves in the value of the stocks:

$$w_{i,t} = w_{i,t-1}\left(1 + r_{i,t-1}\right) \qquad (4)$$

Daily returns are calculated as:

$$r_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \times 100\% \qquad (5)$$

As transaction costs play a vital role in a trading strategy's effectiveness (Timmermann and Granger, 2004; Do and Faff, 2012), we generate net return series by subtracting all the explicit and implicit transaction costs, namely, commissions, taxes, short-selling fees, and market impact at each trade. Given that the transaction costs vary to different investors and fluctuate over time, we set an average transaction costs of 1% per trade [14] throughout the sample which is very conservative and alleviates the concern for

---

[13] It means that every day closing prices are used to determine whether a pair should be opened. If a signal is received, one is assumed to be able to buy the very second the stocks for the same closing prices (that was used to determine the signal). A bit theoretical, but almost doable even in practice.

[14] The commissions and taxes for mainland China investors to trade Hong Kong stocks are less than 0.3% per trade, the daily short selling cost is about 0.02% (sources: http://edu.sse.com.cn/col/shhkconnect/faq/jyyjs/ and http://www.htsc.com.cn/browser/rzrq/index.jsp#).

nonsynchronous trading. In literature, Gatev *et al.* (2006), Do and Faff (2010) and Bowen and Hutchinson (2016) estimate the profitability of pairs trading before transaction costs. However, these papers consider the effects of the bid-ask spread and 1-day delay of trading, which are noisy proxies for market impact. Do and Faff (2012) estimates an average one-way cost (commission plus market impact) of 0.6% for the full sample and time-varying costs for subsamples. Broussard and Vaihekoski (2012) set the transaction costs at 0.2% per trade throughout the sample. Bowen *et al.* (2010) using a range of transaction costs from 0 to 0.15%, and find that increasing transaction costs from 0 to 0.15% can reducing excess returns of the strategy from 15.2% to 7.0%.

The net return on a pair, $R_{p,t}$, can also be considered as the net excess return, given that the risk-free rate in Equation (3) is cancelled if we calculate the excess return on both legs of the pair. The pairs trading strategy is designed to begin without any investment, but as the weights ($w_{L,t}$ and $w_{S,t}$) change in Equation (3), the position has either a positive or negative net value over the trading period.[15]

Based on the return series for each pair, we can generate the net excess return for a portfolio of the pairs at time t ($R_{P,t}$). We employ two alternative weighting schemes. The first one is named Committed Capital scheme, which basically commits equal amounts of investment to each one of the N pairs. If the pair is not opened, or closed during the trading period, the investment is still committed to the pair. We assume zero return for non-open pairs. The second one is named Fully Invested scheme, which assumes investment is always divided between the pairs that are open. Thereby, the portfolio return based on Committed Capital scheme is more conservative than the Fully Invested scheme as it assumes money is always divided between the N pairs rather than the pairs that are open, and N is always no less than the number of pairs that are open.

Following Broussard and Vaihekoski (2012), we use two different approaches to changes in the weights in both schemes, namely, the Equally-Weighted approach and the Value-Weighted approach. The Equally-Weighted approach uses the sum of returns divided either by the number of pairs (N) or the number of pairs that are open at any given date for the Committed Capital scheme and Fully Invested scheme, respectively. The Value-Weighted Committed Capital approach for a portfolio of N pairs is calculated as shown:

$$R_{P,t} = \frac{\sum_{p \in N} W_{p,t} \times R_{p,t}}{\sum_{p \in N} W_{p,t}} \tag{6}$$

The weight for each pair ($W_{p,t}$) is calculated as $W_{p,t} = W_{p,t-1}(1 + R_{p,t-1})$, and the weight ($W_{p,t}$) with a starting value of 1. If a pair is not opened, the return is 0, and the weight does not change. For Value-Weighted Fully Invested approach, we just need to replace N in Equation (6) by the number of pairs that are open. This means that for the Fully Invested weighting scheme, the investment from a closed pair is invested in the other pairs that are open. If the pair is re-opened, the investment is invested back by moving the capital between the pairs according to their relative weights. Only stocks that are alive throughout the formation period and trading period are considered in the investigation. Daily returns are compounded to generate monthly returns. By repeating the 12×6 implementation cycle forward 1-month each time, there are 6 overlapping trading periods of excess returns, which we average to yield monthly net excess return series for the pairs trading.[16]

---

[15] Technically, it is easy to force the weights to remain the same implying a net 0 position, however, in practice it means daily rebalancing the positions and would leading to high transaction costs.

[16] Broussard and Vaihekoski (2012) rolls the 12×6 implementation cycle forward every 6-month and the pairs are formed using data either from January to December within a calendar year, or from July to June the following year. They calculate the monthly returns using daily cumulative return series.

3.3 Performance evaluation

The profitability of the strategy is estimated as the regression intercept for the monthly net excess return series for a portfolio of pairs against Fama and French (2015) 5-factor plus momentum and short-term reversal factors as shown:

$$R_{P,t} = \alpha + \beta_{R_M}R_{M,t} + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{RMW}RMW_t + \beta_{CMA}CMA_t \qquad (7)$$
$$+ \beta_{MOM}MOM_t + \beta_{REV}REV_t + \varepsilon_t$$

Where, $R_{P,t}$ is the monthly net excess return for a portfolio of N pairs at time t; α is the intercept, or measure of abnormal net return; $R_{M,t}$ is the market factor, defined as market portfolio's return subtracts the risk-free rate at time t; SMB_t, HML_t, RMW_t, CMA_t, MOM_t and REV_t are the size factor, value factor, profitability factor, investment factor, momentum factor, and short-term reversal factor[17], respectively. β_{R_M}, β_{HML}, β_{RMW}, β_{CMA}, β_{MOM}, β_{REV} are the regression coefficients. ε_t is the regression residual at time t. A statistically significant and positive (negative) α indicates the portfolio is statistically and economically profitable (unprofitable) after adjusting for the relevant risk factors and transaction costs, and thereby suggest the market is inefficient (efficient). A statistically insignificant β_{R_M} would indicate the portfolio return is insensitive to market return, and thereby market-neutral. See Table A1 for a detailed summary of the factors.

4. Empirical Results

4.1 Is pairs trading profitable in mainland China and Hong Kong?

To investigate whether the integration of stock markets matter, we first present the profitability of the strategy when trading is isolated to just mainland China (Panel A) and then just Hong Kong (Panel B). We then display the profitability of pairs trading across mainland China and Hong Kong (Panels C and D), in Tables 1 through 3.

Table 1 displays descriptive statistics for the monthly net excess returns for a portfolio of Top 5 pairs on different weighting schemes over the full sample period. Panel A shows the results for China Security Index 300 (CSI300) constituents, which consist of the 282 large- and mid cap stocks listed on A-shares in mainland China. Under the scenario that without trading delay, the average monthly net excess returns are positive but insignificant at conventional significance levels across different weighting schemes. All monthly excess returns show positive skewness and kurtosis higher than 3, indicating the returns skewed to the right and produce higher and more extreme outliers than the normal distribution. Jarque-Bera tests clearly reject the null hypothesis that the monthly net returns are normally distributed at the 1% significance level. The monthly net returns range from -9.17% (equivalent to -68.47% per annual) on the Value-Weighted Fully Invested scheme to 13.99% (equivalent to 381.29% per annual) on the Equally-Weighed Fully Invested scheme. The abnormal returns (α) and coefficients for market factor (β_{R_M}) are all statistically insignificant at conventional significance levels, suggesting that pairs trading is unprofitable on CSI300 constituents albeit the strategy is market-neutral.[18] Under the trading scenario with 1-day delay, the average monthly net excess returns are significant at 1% significance level with values ranging from -1.30% (equivalent to -14.53% per annual) on the Value-Weighted Fully Invested scheme to -0.6% (equivalent to -6.97% per annual) on the Value-Weighted Committed Capital scheme. The abnormal returns are significantly negative at 1% significance level with values between -1.30% (equivalent to -14.53% annually) on the Fully Invested schemes and -0.67% (equivalent to 7.75% annually) on the Value-Weighed Committed Capital scheme, suggesting that the pairs trading strategy statistically and economically lose money. However, the monthly net excess returns remain non-normally distributed and market-neutral. The poor performance of pairs

---

[17] The authors construct the Chinese version of risk factors by following the methodology of Gatev *et al.* (2006) and Fama and French (2015). All factors are available upon request.

[18] For the sake of brevity, we do not present the coefficients for the rest of risk factors but the results are all available upon request.

trading in Panel A is probably due to our dataset is constrained to highly-liquid large-cap stocks, which are more likely to be efficiently priced. Moreover, our investigation is based on a very conservative estimation of trading costs. Do and Faff (2012) also find that after adjusting for commissions, market impact, and short-selling fees, the original trading stratgies of Gatev *et al.* (2006) loses its profitability.

Panel B shows the results for Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents, which include 261 H-shares and other stocks listed on the Stock Exchange of Hong Kong. With a few exceptions, such as the average excess returns are negative but insignificant under the trading scenario of no trading delay, and therefore the findings in Panel B are consistent with Panel A.

Panel C shows the results for Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents, which include 126 stocks for 63 distinct firms dual-listed on A-shares in mainland China and H-shares in Hong Kong. Under the no delay trading scenario, the average monthly net excess returns are statistically significant and positive at 10% significance level on the Fully Invested weighting schemes but not for Committed Capital schemes. Furthermore, the abnormal return is significant at the 10% significance level on the Value-Weighted Fully Invested scheme with a magnitude of 0.53% (equivalent to 6.55% per annum). Under the 1-day delay trading scenario, the magnitudes of average excess returns and abnormal returns are less than the corresponding values in Panels A and B, although the values are still significantly negative, suggesting HSAHP constituents lose less compare to CSI300 constituents and HSHKI constituents. The previous Chinese literature (Li *et al.*, 2014) also find pairs trading profitable on dual-listed stocks by using cointegration method, however, our results are not directly comparable as we use very different methodology and dataset. The outperformance of Panel C is probably due to the fundamental values of HSAHP constituents are less uncertain, given that dual-listed firms share the same cash-flow source (Broussard and Vaihekoski, 2012). The historically low integration of mainland China and Hong Kong stock markets is another possible reason prevented arbitrageurs from taking the advantage of disequilibrium stock prices immediately and thereby makes pairs trading profitable (Schotman and Zalewska, 2006; Jacobs and Weber, 2015).

Panel D shows the results for all constituents where investors can trade across all stock exchanges, which comprised of the first 3 datasets and contain 622 unique stocks. Under the no delay trading scenario, the average monthly net excess returns are significant at the 5% level with values above 0.30% on different weighting schemes. Abnormal returns are all significant at the 5% level with values ranging between 0.39% (equivalent to 4.78% per annual) on the Equally-Weighed Committed Capital scheme and 0.72% (equivalent to 8.99% per annual) on the Value-Weighed Fully Invested scheme. Panel D does suggests that the trading strategy including all the stocks across mainland China and Hong Kong is more profitable than strategy including only companies have dual listing. The rest of findings in Panel D are similar to Panels A through C. The magnitudes of annualized abnormal returns are much lower than what Broussard and Vaihekoski (2012) found in Finland (up to 10.82% per annual).[19]

Overall, Table 1 suggests that pairs trading in the mainland China and the Hong Kong stock market is unprofitable and therefore both markets are efficient over the full sample period. However the two markets are not well integrated, as there are profitable arbitrage opportunities across the two markets. It seems it is the integration rather than the population of stocks that plays a more important role in determining the efficiency of market, given that the pairs trading is profitable on HSAHP constituents (126 stocks) and ALL constituents (622 stocks) but unprofitable on CSI300 constituents (282 stocks) and HSHKI large- and mid-cap constituents (261 stocks). The findings in Table 1 are qualitatively consistent with Jacobs and Weber (2015) that the strategy is quite successful for pairs that are hard to arbitrage or less visible.

Table 1 about here

To examine whether the number of pairs matters to profitability, Tables 2, 3 and A3 present the descriptive statistics for a portfolio of Top 10, Top 20 and all pairs, respectively. With a few exceptions, the findings

---

[19] Broussard and Vaihekoski (2012) estimate the abnormal return using Capital Asset Pricing Model under the scenario of 1-day trading only.

in Tables 2, 3 and A3 are consistent with Table 1. The trading strategy of including all pairs is not competitive relative to the strategy of Top N pairs since it requires a substantial amount of transaction costs and execution time which, in turn, reduces the profitability. Thereby, we do not present the empirical results of all pairs in the subsequent analysis. But all the results are available upon request. Specially, Tables 1 through 3 suggest the HSAHP constituents show the best performance for a portfolio of Top 20 pairs. On the contrary, ALL constituents report the best profitability for a portfolio of Top 5 pairs but worst and unprofitable performance for a portfolio of Top 20 pairs.

<div align="center">Tables 2 and 3 about here</div>

To explore the characteristics of pairs trading, Table 4 presents the trading statistics for a portfolio of Top 5, 10 and 20 pairs, respectively. For CSI300 constituents, the average price deviations trigger for opening pairs[20] are 0.12, 0.13 and 0.14 for a portfolio of Top 5, 10, 20 pairs, respectively. The average number of times a pair is opened per 6-month period increases from 7.49 for Top 5 pairs, to 14.08 for Top 10 pairs, and 26.42 for Top 20 pairs. In a round-trip style, each pair is opened and closed about 1.50 times for Top 5 pairs, 1.41 times for Top 10 pairs, and 1.32 times for Top 20 pairs. Under the scenario of no trading delay, on average, each pair is held open about 51.57 days, 52.78 days and 53.37 days for the Top 5, 10 and 20 pairs portfolio, respectively. Under the trading scenario of 1-day delay, the average opening time for each pair is about 2 days less than their no delay trading scenario counterparts. The similar patterns held for the other 3 datasets. It seems the frequency of transactions needed to implement the pairs trading is very low and the average holding period per pair is quite long which, in turn, induce stability of the pairs trading strategy. Literature (Gatev *et al.*, 2006; Broussard and Vaihekoski, 2012; Bowen and Hutchinson, 2016) report the similar findings.

<div align="center">Table 4 about here</div>

Figure 1 presents the cumulative monthly net excess returns for a portfolio of Top 5 pairs and the market returns in excess of risk-free rate. Under the no delay trading scenario, the portfolio returns on Fully Invested schemes outperform the market on HSAHP constituents and ALL constituents by the end of the sample. CSI300 constituents and HSHKI large- and mid-cap constituents beat the market sometimes during the sample. The results under the 1-day delay trading scenario are not reported as not one portfolio of pairs beats the market.[21]

<div align="center">Figure 1 about here</div>

The findings so far consistently suggest the integration of financial markets and number of pairs matter to pairs trading. Specially, the dual-listed stock class environment and historically isolated stock markets in China provides an opportunity to arbitrage for pairs trading, even taking the transaction costs into account, which favour the hypothesis of Broussard and Vaihekoski (2012) and the arguments of the AMH. However, the optimal number of pairs remains unclear. Moreover, we caution the non-professional retail investors to execute the pairs trading as the profitability is quite sensitive to the speed of trading, suggesting the disequilibrium of stock prices and arbitrage opportunity disappear very quick.

4.2 Is the profitability time-varying?

To investigate whether the profitability of pairs trading is time-varying, Figure 2 displays the t-statistics of abnormal returns (t_α) for a portfolio of Top 5 pairs using Equation (7) with 120-month window length rolling forward 1-month each time. Broadly speaking, the t-statistics on different weighting schemes follow the similar pattern. Under the no delay trading scenario, the t-statistics for CSI300 constituents decline from 0.50 in 12/2006 to -0.50 by 12/2007; then rebound; stay around 2 from 12/2009 to 03/2012, indicating pairs trading for CSI300 is significantly profitable at 5% significance level and thereafter, the t-statistics move toward 0 by 07/2017. For HSHKI large- and mid-cap constituents, the t-statistics move between -

---

[20] The calculation is based on normalized stock price series.
[21] However the 1-day delay trading results are available upon request from the corresponding author.

1.60 and 1.50 throughout the sample, suggesting persistently unprofitable. For HSAHP constituents, the Fully Invested scheme based *t*-statistics decline from 2.0 in 12/2006 to -1.60 in 03/2012, and then rebound to about 0.50 by 07/2017; the Committed Capital based *t*-statistics move between -2.20 and 1.20, and fluctuate around -2.0 from 12/2012 to 02/2014, indicating significantly beaten by the market. The *t*-statistics for ALL constituents decline with a downward trend from 4.0 in 12/2006 to 2.0 in 12/2011 and then move toward 0 by 07/2017, suggesting the profitability disappears since 2012. Under the 1-day delay trading scenario, the *t*-statistics of abnormal returns fluctuate around -2 and remain less than 0 throughout the sample for all the trading strategies, suggesting persistently unprofitable and significantly lose money during some of the subsamples. Overall, Figure 2 suggests the profitability of pairs trading is time-varying and the profitability disappears in recent years.

Figure 2 about here

As a cross comparison, Table 5 shows the descriptive statistics for the percentage of subsamples with statistically significant abnormal return at 5% significance level for a portfolio of Top 5, 10 and 20 pairs, respectively. Under the no delay trading scenario, Panel A suggests CSI300 constituents generate significant abnormal returns in at least 7.75% (10 out of 129) of subsamples for a portfolio of Top 5 pairs, and the ratio increases up to 26.36% for a portfolio of Top 20 pairs. In Panel B, HSHKI large- and mid-cap constituents fail to yield any significant positive abnormal return. Panel C suggests the profitability on HSAHP constituents are very weak for a portfolio of Top 5 pairs but quite strong for a portfolio of Top 20 pairs. Panel D suggests that the strategy is profitable on ALL constituents in at least 21.71% of subsamples for a portfolio of Top 20 pairs, and the ratio jumps up to 55.04% for a portfolio of Top 10 pairs.

Overall, Figure 2 and Table 5 suggest the Hong Kong stock market is far more efficient than the mainland China market historically. However, the efficiency in mainland China stock market is improving in recent years. There is no evidence to indicate the efficiency of Hong Kong market deteriorates over time. In terms of timing, the openness of China stock market to froeign investors since 2003 constituents an exogenous shock to the arbitrageurs competition to take advantage of pairs trading profitability. Take the QFII, RQFII and mutual market access shcemes into account, the total amount of foreign capital accounts for about 4% of A-shares market value in 2002 and 7% by 2016.[22]

Table 5 about here

4.3 Is the time-varying profitability due to data-mining?

We apply a modified version of Fama and French (2010) bootstrap simulation to investigate whether the time-varying profitability in previous section is due to data-mining. The original Fama and French (2010) bootstrap simulation applies to a large number of the US mutual funds throughout a given sample period, and has been widely used in mutual funds literature (Gallefoss *et al.*, 2015; Blake *et al.*, 2017). We apply the bootstrap simulation to a given return time-series over a number of subsamples.

Specially, we initially estimate Equation (7) using rolling-window regression with 120-month window size and rolling forward 1-month each time, as shown in Figure 2. For each subsample s, we save the coefficient estimates $\left\{\widehat{\alpha_s}, \widehat{\beta_{s,R_M}}, \widehat{\beta_{s,SMB}}, \widehat{\beta_{s,HML}}, \widehat{\beta_{s,RMW}}, \widehat{\beta_{s,CMA}}, \widehat{\beta_{s,MOM}}, \widehat{\beta_{s,REV}}\right\}$, the time-series of estimated residuals $\widehat{\varepsilon_{s,t}} = [\widehat{\varepsilon_{s,1}}, \cdots, \widehat{\varepsilon_{s,120}}]$, and the *t*-statistic of intercept $\widehat{t_{\alpha_s}}$. Then, we generate a pseudo time-series of residuals $\widehat{\varepsilon_{s,t}^b} = [\widehat{\varepsilon_{s,1}^b}, \cdots, \widehat{\varepsilon_{s,120}^b}]$ by randomly collecting residuals $\widehat{\varepsilon_{s,t}^b}$ from the original residual vector $\widehat{\varepsilon_{s,t}}$ with replacement. Similarly, we generate pseudo time-series of risk factors $\left\{\widehat{R_{M,s,t}^b}, \widehat{SMB_{s,t}^b}, \widehat{HML_{s,t}^b}, \widehat{RMW_{s,t}^b}, \widehat{CMA_{s,t}^b}, \widehat{MOM_{s,t}^b}, \widehat{REV_{s,t}^b}\right\}$ by randomly collecting values with replacement from the original risk vector $\{R_{M,s,t}, SMB_{s,t}, HML_{s,t}, RMW_{s,t}, CMA_{s,t}, MOM_{s,t}, REV_{s,t}\}$. b is

the bootstrap index. Next, produce a time-series of pseudo excess returns for this subsample s, imposing the null hypothesis of zero true performance ($\alpha_s = 0$) by using:

$$R^b_{P,s,t} = 0 + \widehat{\beta_{s,R_M}}\widehat{R^b_{M,s,t}} + \widehat{\beta_{s,SMB}}\widehat{SMB^b_{s,t}} + \widehat{\beta_{s,HML}}\widehat{HML^b_{s,t}} + \widehat{\beta_{s,RMW}}\widehat{RMW^b_{s,t}} \qquad (8)$$
$$+ \widehat{\beta_{s,CMA}}\widehat{CMA^b_{s,t}} + \widehat{\beta_{s,MOM}}\widehat{MOM^b_{s,t}} + \widehat{\beta_{s,REV}}\widehat{REV^b_{s,t}} + \widehat{\varepsilon^b_{s,t}}$$

The pseudo excess return $R^b_{P,s,t}$ is then regressed on Equation (7), estimating an intercept $\widetilde{\alpha^b_s}$ and t-statistic of intercept $\widetilde{t^b_{\alpha,s}}$. Repeating the above procedures through all 129 subsamples, $s = 1, \cdots, 129$, we get a draw from the time-series of bootstrap $\widetilde{\alpha^b_s}$ and $\widetilde{t^b_{\alpha,s}}$, respectively. We rank the $\widetilde{\alpha^b_s}$ from the smallest value to biggest value to yield a luck (pure chance or zero-skill) cumulative distribution function of the alphas, and in the same way for $\widetilde{t^b_{\alpha,s}}$. Repeating this for all bootstrap iterations 10,000 times, $b = 1, \cdots, 10,000$. Then, we can compare the luck distribution of simulated $\widetilde{\alpha^b_s}$ to actual $\widehat{\alpha_s}$, or simulated $\widetilde{t^b_{\alpha,s}}$ to actual $\widehat{t_{\alpha_s}}$. Throughout the paper, the standard errors are corrected for autocorrelation (0 lags) and heteroscedasticity through the Newey and West (1987) method for two reasons. First, the majority of rolling windows regressions do not report autocorrelation at the conventional significance level by using the Breusch-Godfrey test. Second, it has the advantage of enhancing comparability between simulated and actual t-statistics through a uniform test specification. This is because the bootstrap simulations consist of random drawings of individual daily returns with replacement, which means the time-series drawn cannot contain any true underlying autocorrelation by design (Meyer *et al.*, 2012).

The rationale behind the bootstrap simulations is to investigate how many subsamples one might anticipate achieves a given level of alpha performance by random chance alone and compares this with the number of subsamples which actually achieved this level of alpha in the 'real world'. For the outperformance (underperformance) subsamples which is measured by alpha, if simulated $\widetilde{\alpha^b_s}$ is greater than actual $\widehat{\alpha_s}$ in less than 5% of the 10,000 simulations, at any given performance order, we reject the null hypothesis that the outperformance (underperformance) of pairs trading is due to good luck (poor stock picking skill) at 95% confidence level and infer that the strategy is genuine (bad luck) (Cuthbertson *et al.*, 2008; Blake *et al.*, 2017).

Table 6 shows results from the bootstrap simulation for a portfolio of Top 5 pairs on CSI300 constituents in Panel A and ALL constituents in Panel B, for the sake of brevity. Additionally, the portfolio returns are based on Equally-Weighed Committed Capital scheme. In Panel A, under the no delay trading scenario, the simulated abnormal returns are higher than the corresponding actual abnormal returns in less than 5% of the simulations from the 1st through the 125th subsamples (the order is ranked by abnormal return), hence reject the null hypothesis that the profitability (unprofitability) periods of pairs trading is due to good luck (poor stock picking ability) at 5% significance level. It implies these poorly performed subsamples, such as the declining profitability of pairs trading in recent years, are simply due to bad luck rather than the poor ability of the strategy. The majority of positive abnormal returns are the result of the strategy's superior stock picking ability. For the 4 best performed subsamples, the simulated abnormal returns are higher than the actual abnormal returns in more than 5% of the 10,000 simulations, suggesting these outperformances are purely good luck. Under the 1-day delay trading scenario, the simulated abnormal returns consistently higher than the actual abnormal returns in more than 95% of the simulations throughout the 129 subsamples, suggesting these poor performances are due to poor ability rather than poor luck which confirms the importance of quick trading. The results using t-statistics of abnormal returns (t_α) are quite similar. The findings in Panel B, on ALL constituents, are consistent with Panel A.

Table 6 about here

4.4 Is the profitability of pairs trading sensitive to market conditions?

Table 7 shows results of Equation (7) estimated across two subsamples, where the market factor is either greater than or equal to zero (short-term up condition), or less than zero (short-term down condition). For

CSI300 constituents, the abnormal returns are significantly negative at 10% significance level when $R_{M,t} <$ 0, indicating the strategy underperforms during short-term market downturn. The coefficients for market factor are all significantly negative at the 5% level, implying the market factor play negative role on the performance of the strategy. None of the rest abnormal returns are statistically significant at conventional significance levels.

Table 7 about here

To examine the performance in longer-term market condition, we re-investigate the analysis in Table 7 by using the prior 12-month market excess returns. Table 8 suggests that the abnormal returns on ALL constituents are significantly positive during longer-term market downturn at 5% significance level, except for a portfolio of Top 20 pairs. All the rest of abnormal returns are insignificant. Tables 7 and 8 suggest the pairs trading strategy in China is somewhat sensitive to the market conditions, which contrast to the UK literature (Bowen and Hutchinson, 2016).

Table 8 about here

4.5 Does the strategy works as a return enhancer?

Given that neither a portfolio of 5 pairs nor 20 pairs provides sufficient diversification, a rational risk-averse investor might combining an active trading portfolio with a well-diversified passive portfolio (Auer and Schuhmacher, 2015). We test whether pairs trading can enhance the performance of a given passive portfolio using the portfolio allocation strategy of Treynor and Black (1973) as shown in Equation (9).

$$w_a = \frac{w_a^*}{1 + (1 - \beta_a)w_a^*}; \text{ with } w_a^* = \frac{\alpha_a/\sigma_{\epsilon_a}^2}{\mu_p/\sigma_p^2} \tag{9}$$

w_a is the weight for a portfolio of pairs, proxy for the active portfolio. α_a and β_a are the intercept and slope in a single-index model of the excess returns of the active portfolio against the excess returns of the passive portfolio, respectively. We use the broad market portfolio to proxy the passive portfolio. $\sigma_{\epsilon_a}^2$ is the variance of the residuals from the single-index model; μ_p and σ_p^2 are the mean and variance of the excess returns for the passive portfolio, respectively.

Table 9 shows the abnormal returns of the combined portfolios under the no delay trading scenario. The HSAHP constituents yield significant positive abnormal returns for a portfolio of Top 10 and 20 pairs on the Fully Invested weighting schemes at 5% significance level. ALL constituents generate significant positive abnormal returns on all weighting schemes for a portfolio of Top 5 and 10 pairs, the value of abnormal returns range from 0.14% (equivalent to 1.69% per annual) to 0.37% (equivalent to 4.53% per annual). All the rest combined portfolios fail to yield significant abnormal returns. Therefore, pairs trading strategy using HSAHP constituents and ALL constituents could be a return enhancer.

Table 9 about here

5. Conclusion

This paper contributes to the literature by providing a comprehensive investigation of the pairs trading on highly-liquid large- and mid-cap stocks listed on mainland China and Hong Kong stock exchanges using a sample period between 1996 and 2017. Over the full sample period, we first find that the pairs trading strategy yields statistically and economically significant net monthly excess returns and net abnormal returns when investors trade across mainland China and Hong Kong stock markets. However, the strategy is unprofitable when investors focus on either mainland China or Hong Kong stock market. The number of pairs in a portfolio matters to the profitability of the strategy. Through estimating rolling-window regression

for the monthly return series against Fama and French (2015) 5-factor plus momentum and short-term reversal factors, we find the profitability of the strategy is time-varying. Specially, the strategy is profitable in mainland China stock market from 2010 to 2012, consistently unprofitable in Hong Kong stock market and profitable across mainland China and Hong Kong stock markets prior to 2011.

By using a modified version of Fama and French (2010) bootstrap simulation, we find the poorly performed subsamples such as the general declining profitability of the strategy in recent years are due to bad luck, the vast majority of profitable subsamples are due to superior stock picking ability and a few of the best performed subsamples are due to good luck. In terms of timing, it seems the integration of Chinese stock markets and permission of short-selling do improve the market efficiency in mainland China but without deteriorating the market efficiency in Hong Kong. Consistent with the literature, we also find the pairs trading shows strong performance during longer-term market turbulence when investors trade across the Chinese stock markets. Finally, we find the pairs trading could be a valuable index enhancer for risk-averse investors. However, we caution the non-professional retail investors to execute pairs trading in China as the strategy is quite sensitive to the speed of execution.

Overall, our findings are consistent with the AMH by showing that the market efficiency is time-varying and that the time-varying performance is not due to data-mining but changes in market conditions. It is the integration of financial markets and market conditions rather than the number of stocks play more important role in driving market efficiency.

Table 1. Monthly net excess returns of pairs trading for a portfolio of top 5 pairs

	No delay				1-day delay			
	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$
Panel A: China Security Index 300 (CSI300) constituents								
Average excess return	0.06	0.07	0.05	0.07	-0.68***	-1.20***	-0.60***	-1.30***
	(0.16)	(0.25)	(0.15)	(0.25)	(0.17)	(0.28)	(0.15)	(0.26)
Standard deviation	1.97	3.22	1.89	3.19	2.22	3.80	2.02	3.51
Skewness	0.11	0.51	0.32	0.58	-0.43	-0.13	-0.19	-0.24
Kurtosis	4.03	4.72	4.39	4.92	4.30	4.43	4.40	5.11
Jarque-Bera statistic	11.5***	4.90***	24.1***	51.9***	25.10***	21.70***	21.70***	48.10***
Minimum	-5.98	-8.33	-5.43	-9.17	-7.53	-12.89	-7.50	-15.43
Maximum	8.01	13.99	8.40	13.48	7.73	11.24	7.24	11.06
Abnormal return	0.02	0.07	0.02	0.05	-0.78***	-1.30***	-0.67***	-1.30***
	(0.18)	(0.28)	(0.17)	(0.27)	(0.21)	(0.32)	(0.17)	(0.28)
β_{R_M}	0.00	-0.01	0.00	-0.01	0.01	-0.00	0.01	-0.01
	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.03)
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents								
Average excess return	-0.15	-0.34	-0.12	-0.30	-0.94***	-1.90***	-0.76***	-1.30***
	(0.19)	(0.35)	(0.17)	(0.34)	(0.23)	(0.44)	(0.19)	(0.30)
Standard deviation	2.33	4.85	2.14	4.71	2.90	5.56	2.46	4.63
Skewness	-0.92	-2.66	-0.54	-2.58	-2.34	-2.77	0.87	-1.32
Kurtosis	6.47	21.98	6.76	22.42	14.25	16.45	15.17	15.45
Jarque-Bera statistic	159***	4,014***	158***	4,174***	1,534***	2,186***	1,562***	1,673***
Minimum	-1.99	-37.36	-1.65	-36.38	-19.25	-36.75	-9.60	-34.18
Maximum	7.09	15.92	6.91	15.47	7.33	11.85	16.62	19.43
Abnormal return	-0.11	-0.35	-0.13	-0.37	-0.82***	-1.70***	-0.87***	-1.40***
	(0.20)	(0.40)	(0.19)	(0.39)	(0.21)	(0.41)	(0.21)	(0.35)
β_{R_M}	0.01	0.03	0.01	0.03	0.01	0.05	0.03	0.04
	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents								
Average excess return	0.12	0.52*	0.13	0.53*	-0.53**	-0.66*	-0.44**	-0.59**
	(0.16)	(0.28)	(0.16)	(0.28)	(0.22)	(0.34)	(0.19)	(0.30)
Standard deviation	1.68	2.91	1.64	2.90	2.34	3.72	2.05	3.29
Skewness	-0.64	0.23	-0.65	0.25	-2.46	-1.78	-2.08	-0.75
Kurtosis	4.83	3.83	5.22	3.82	12.53	11.51	11.49	5.23
Jarque-Bera statistic	51.8***	9.25***	68.6***	9.42***	1,189***	879***	925***	74.70***
Minimum	-6.49	-9.31	-6.33	-9.44	-14.95	-25.24	-11.78	-14.67
Maximum	4.69	1.51	4.55	1.10	3.98	7.54	4.16	7.69
Abnormal return	0.10	0.51	0.12	0.53*	-0.6**	-0.8**	-0.49**	-0.71**
	(0.18)	(0.32)	(0.18)	(0.31)	(0.24)	(0.37)	(0.20)	(0.32)
β_{R_M}	-0.00	-0.00	0.01	0.00	-0.03	-0.05	-0.01	-0.02
	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.06)	(0.02)	(0.04)
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)								
Average excess return	0.31**	0.57**	0.33***	0.62***	-0.48***	-0.75***	-0.38***	-0.74***
	(0.13)	(0.23)	(0.12)	(0.22)	(0.15)	(0.28)	(0.13)	(0.24)
Standard deviation	1.85	3.09	1.75	2.96	2.16	4.04	1.93	3.25
Skewness	-0.23	0.46	0.05	0.61	-1.03	-0.12	-0.13	0.16
Kurtosis	3.71	4.58	3.28	4.20	6.77	12.95	4.34	5.0
Jarque-Bera statistic	7.41**	34.3***	0.93	3.60***	191***	1,024***	19.10***	42.30***
Minimum	-6.03	-8.78	-4.62	-6.28	-1.23	-24.92	-6.41	-11.68
Maximum	5.18	12.27	4.86	11.37	5.15	23.04	5.27	11.50
Abnormal return	0.39**	0.69**	0.40***	0.72***	-0.43**	-0.71**	-0.35**	-0.73***
	(0.16)	(0.28)	(0.15)	(0.26)	(0.17)	(0.32)	(0.15)	(0.27)
β_{R_M}	-0.01	-0.03	-0.00	-0.02	-0.01	-0.01	-0.00	-0.01
	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)

Notes: This table displays results of compounded monthly return series from 01/1997 to 07/2017. $R_P\_EW\_CC$ is the portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is the Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is the Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is the Value-Weighted Fully Invested scheme. The null hypothesis for the Jarque-Bera test is that the variable is normally distributed. The row labelled "Abnormal return" presents the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row labelled "β_{R_M}" is the coefficient for the market factor. The authors construct the Chinese-version of risk factors by following Gatev *et al.* (2006) and Fama and French (2015). The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. The returns are in % term.

Table 2. Monthly net excess returns of pairs trading for a portfolio of top 10 pairs

	No delay				1-day delay			
	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$
Panel A: China Security Index 300 (CSI300) constituents								
Average excess return	0.06	0.14	0.05	0.13	-0.65***	-1.10***	-0.60***	-1.10***
	(0.11)	(0.19)	(0.1)	(0.19)	(0.13)	(0.22)	(0.12)	(0.20)
Standard deviation	1.33	2.38	1.29	2.37	1.61	2.85	1.44	2.64
Skewness	0.12	0.66	0.2	0.83	-0.14	0.62	-0.08	0.56
Kurtosis	3.46	4.48	3.62	5.09	3.9	5.14	4.40	5.61
Jarque-Bera statistic	2.68	4.3***	5.66*	73***	9.14***	62.8***	2.5***	83***
Minimum	-3.30	-5.19	-3.65	-5.07	-5.44	-9.18	-6.39	-9.04
Maximum	4.34	9.12	4.07	9.73	5.15	2.0	4.30	1.56
Abnormal return	0.02	0.11	0.01	0.08	-0.75***	-1.20***	-0.68***	-1.30***
	(0.12)	(0.21)	(0.11)	(0.20)	(0.16)	(0.24)	(0.13)	(0.21)
β_{R_M}	0.01	-0.01	0.01	-0.00	0.01	0.02	0.01	0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents								
Average excess return	0.22	0.23	0.83	0.84	-0.65**	-1.13**	0.21	-0.41
	(0.36)	(0.49)	(0.79)	(0.9)	(0.31)	(0.45)	(0.80)	(0.69)
Standard deviation	4.23	5.54	9.40	1.63	3.59	5.11	9.42	7.82
Skewness	6.52	3.94	9.59	8.61	4.07	2.29	9.04	7.22
Kurtosis	57.74	32.18	102.7	89.97	36.20	21.48	87.83	67.44
Jarque-Bera statistic	334***	9,438***	1,100***	8,100***	1,200***	3,746***	7,800***	4,500***
Minimum	-7.52	-17.12	-7.69	-18.52	-13.85	-22.01	-7.14	-13.13
Maximum	41.12	45.01	114.15	124.85	28.99	33.28	101.73	81.87
Abnormal return	0.55	0.55	1.34	1.33	-0.33	-0.74	0.65	-0.10
	(0.52)	(0.65)	(1.16)	(1.29)	(0.40)	(0.53)	(1.14)	(0.93)
β_{R_M}	-0.04	-0.05	-0.06	-0.06	-0.04	-0.05	-0.05	-0.05
	(0.03)	(0.04)	(0.05)	(0.06)	(0.02)	(0.04)	(0.05)	(0.05)
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents								
Average excess return	0.16	0.68***	0.15	0.67***	-0.34**	-0.36	-0.28**	-0.32
	(0.11)	(0.26)	(0.10)	(0.25)	(0.14)	(0.26)	(0.12)	(0.24)
Standard deviation	1.14	2.49	1.11	2.48	1.50	2.74	1.34	2.55
Skewness	-0.44	0.72	-0.37	0.75	-1.92	-0.38	-1.51	0.12
Kurtosis	5.93	4.38	5.95	4.32	1.84	5.30	9.92	3.90
Jarque-Bera statistic	96.3***	41.30***	95.60***	4.90***	788***	6.50***	589***	8.83**
Minimum	-4.59	-5.59	-4.36	-5.51	-8.44	-11.70	-7.33	-7.68
Maximum	4.70	1.51	4.59	1.10	3.72	7.70	3.73	7.60
Abnormal return	0.14	0.70**	0.14	0.68**	-0.37**	-0.43	-0.30**	-0.40
	(0.11)	(0.28)	(0.11)	(0.28)	(.15)	(.28)	(0.13)	(0.25)
β_{R_M}	0.01	0.02	0.02	0.02	0.00	0.01	0.02	0.03
	(0.01)	(0.03)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.02)
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)								
Average excess return	0.23**	0.44**	0.26***	0.46***	-0.51***	-0.92***	-0.41***	-0.88***
	(0.10)	(0.17)	(0.09)	(0.17)	(0.11)	(0.20)	(0.10)	(0.18)
Standard deviation	1.18	2.18	1.13	2.12	1.36	2.47	1.16	2.20
Skewness	-0.10	0.90	0.16	0.99	-1.06	-0.01	-0.32	0.45
Kurtosis	3.55	6.51	3.31	6.19	7.20	9.05	4.50	8.65
Jarque-Bera statistic	3.45	161***	2.05	146***	228***	378***	27.60***	338***
Minimum	-4.09	-6.07	-3.08	-5.02	-8.12	-14.07	-5.53	-11.74
Maximum	3.58	11.83	3.88	11.18	3.39	12.11	2.95	11.50
Abnormal return	0.27**	0.53**	0.31***	0.57***	-0.50***	-0.84***	-0.38***	-0.82***
	(0.11)	(0.21)	(0.11)	(0.20)	(0.14)	(0.24)	(0.11)	(0.22)
β_{R_M}	-0.00	-0.01	0.00	0.00	-0.01	-0.01	-0.00	-0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)

Notes: This table displays results of compounded monthly return series from 01/1997 to 07/2017. $R_P\_EW\_CC$ is the portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is the Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is the Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is the Value-Weighted Fully Invested scheme. The null hypothesis for the Jarque-Bera test is that the variable is normally distributed. The row labelled "Abnormal return" presents the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row labelled "β_{R_M}" is the coefficient for the market factor. The authors construct the Chinese-version of risk factors by following Gatev *et al.* (2006) and Fama and French (2015). The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. The returns are in % term.

Table 3. Monthly net excess returns of pairs trading for a portfolio of top 20 pairs

	No delay				1-day delay			
	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$
Panel A: China Security Index 300 (CSI300) constituents								
Average excess return	0.09	0.19	0.11	0.21	-0.58***	-1.0***	-0.52***	-1.0***
	(0.09)	(0.16)	(0.09)	(0.16)	(0.11)	(0.19)	(0.09)	(0.17)
Standard deviation	1.06	2.04	1.02	2.03	1.26	2.38	1.11	2.18
Skewness	0.21	0.68	0.66	0.84	-0.53	0.32	-0.15	0.46
Kurtosis	4.40	5.18	5.34	5.36	5.79	6.06	5.40	6.80
Jarque-Bera statistic	22.1***	67.6***	74.50***	86.20***	92***	101***	6.30***	158***
Minimum	-3.36	-4.92	-2.92	-5.04	-5.15	-9.18	-4.10	-9.04
Maximum	4.48	9.12	4.96	9.18	5.04	1.41	4.71	1.56
Abnormal return	0.08	0.22	0.11	0.24	-0.61***	-1.10***	-0.53***	-1.10***
	(0.10)	(0.17)	(0.09)	(0.17)	(0.12)	(0.20)	(0.10)	(0.17)
β_{R_M}	-0.00	-0.02	-0.00	-0.02	0.00	0.01	-0.00	0.01
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents								
Average excess return	0.13	-0.07	0.78	0.57	-0.73***	-1.50***	0.17	-1.40***
	(0.24)	(0.37)	(0.75)	(0.84)	(0.20)	(0.35)	(0.76)	(0.29)
Standard deviation	2.34	3.74	7.67	8.51	2.08	3.67	8.20	3.39
Skewness	4.44	0.74	8.72	7.20	2.35	0.08	7.88	2.72
Kurtosis	33.95	12.34	89.2	69.32	22.62	1.43	66.87	3.11
Jarque-Bera statistic	110***	923***	8,000***	4,800***	4,206***	569***	4,500***	7,805***
Minimum	-4.56	-17.12	-4.58	-18.52	-6.86	-16.77	-6.19	-13.13
Maximum	18.64	21.71	9.75	94	15.64	2.59	8.03	29.14
Abnormal return	0.30	0.13	1.14	0.89	-0.55**	-1.20***	0.34	-1.20***
	(0.35)	(0.45)	(1.06)	(1.16)	(0.26)	(0.35)	(0.94)	(0.37)
β_{R_M}	-0.04*	-0.05	-0.07	-0.08	-0.02	-0.03	-0.06	-0.03
	(0.02)	(0.03)	(0.05)	(0.06)	(0.02)	(0.03)	(0.05)	(0.03)
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents								
Average excess return	0.11	0.72***	0.11	0.72***	-0.28***	-0.39	-0.19*	-0.27
	(0.09)	(0.26)	(0.08)	(0.25)	(0.10)	(0.25)	(0.11)	(0.26)
Standard deviation	0.86	2.47	0.84	2.48	1.09	2.61	1.45	2.95
Skewness	-0.23	0.67	-0.09	0.73	-1.32	0.03	6.76	2.48
Kurtosis	4.65	4.0	4.60	4.0	9.09	3.63	87.83	22.56
Jarque-Bera statistic	3.30***	29.1***	26.7***	32.2***	455***	4.09	7,600***	4,209***
Minimum	-2.65	-5.37	-2.48	-5.27	-5.65	-8.87	-4.86	-8.17
Maximum	2.80	1.51	2.76	1.10	4.19	7.08	17.35	24.67
Abnormal return	0.10	0.74***	0.10	0.74***	-0.29***	-0.44*	-0.21**	-0.36
	(0.08)	(0.28)	(0.07)	(0.28)	(0.10)	(0.26)	(0.10)	(0.25)
β_{R_M}	0.01	0.01	0.01	0.02	0.00	0.00	0.02	0.03
	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)	(0.01)	(0.03)
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)								
Average excess return	0.17*	0.27*	0.17**	0.27*	-0.57***	-1.0***	-0.50***	-0.98***
	(0.10)	(0.17)	(0.08)	(0.16)	(0.11)	(0.19)	(0.09)	(0.17)
Standard deviation	0.98	1.92	0.93	1.85	1.19	2.14	0.99	1.92
Skewness	-0.21	0.97	0.09	1.04	-1.32	0.60	-0.53	1.03
Kurtosis	4.75	8.52	4.09	8.01	8.59	9.54	4.92	1.19
Jarque-Bera statistic	33.6***	354***	12.6***	304***	396***	457***	49.40***	578***
Minimum	-4.47	-5.73	-3.62	-4.91	-7.79	-9.56	-5.19	-7.44
Maximum	2.76	11.83	2.64	11.18	2.05	12.11	1.93	11.50
Abnormal return	0.13	0.30	0.14	0.31*	-0.64***	-1.0***	-0.54***	-0.99***
	(0.10)	(0.20)	(0.09)	(0.19)	(0.14)	(0.25)	(0.11)	(0.22)
β_{R_M}	0.00	-0.01	0.00	-0.01	-0.00	-0.01	0.01	0.00
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)

Notes: This table displays results of compounded monthly return series from 01/1997 to 07/2017. $R_P\_EW\_CC$ is the portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is the Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is the Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is the Value-Weighted Fully Invested scheme. The null hypothesis for the Jarque-Bera test is that the variable is normally distributed. The row labelled "Abnormal return" presents the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row labelled "β_{R_M}" is the coefficient for the market factor. The authors construct the Chinese-version of risk factors by following Gatev *et al.* (2006) and Fama and French (2015). The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. The returns are in % term.

Table 4. Trading statistics

	Top 5 pairs	Top 10 pairs	Top 20 pairs
Panel A: China Security Index 300 (CSI300) constituents			
Average price deviation trigger for opening pairs	0.12	0.13	0.14
Average number of times a pair is opened per 6-month period	7.49	14.08	26.42
Average number of round-trip trades per pair	1.50	1.41	1.32
Average time pairs are open in trading days (No delay)	51.57	52.78	53.37
Average time pairs are open in trading days (1-day delay)	49.75	50.97	51.52
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents			
Average price deviation trigger for opening pairs	0.13	0.14	0.16
Average number of times a pair is opened per 6-month period	6.62	13.27	25.13
Average number of round-trip trades per pair	1.32	1.33	1.26
Average time pairs are open in trading days (No delay)	52.62	53.63	55.05
Average time pairs are open in trading days (1-day delay)	50.81	52.04	53.64
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents			
Average price deviation trigger for opening pairs	0.12	0.14	0.16
Average number of times a pair is opened per 6-month period	5.68	10.02	25.13
Average number of round-trip trades per pair	1.14	1.00	1.26
Average time pairs are open in trading days (No delay)	53.84	54.15	55.05
Average time pairs are open in trading days (1-day delay)	51.31	52.50	53.64
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)			
Average price deviation trigger for opening pairs	0.10	0.11	0.12
Average number of times a pair is opened per 6-month period	7.44	14.52	27.83
Average number of round-trip trades per pair	1.49	1.45	1.39
Average time pairs are open in trading days (No delay)	46.55	47.81	49.81
Average time pairs are open in trading days (1-day delay)	44.78	46.10	48.10

Notes: Average price deviation trigger for opening a pair is calculated as the price difference divided by the average of the prices. Average number of times a pair is opened per 6-month period is the average number of pairs which are opened at least once in the trading period. Average number of round-trip trades per pair is the average number of times a pair opens and closes in each trading period. Average time pairs are open in trading days is the number of days a pair is open on average.

Table 5. Percentage of signficiant subsamples

	Top 5 pairs		Top 10 pairs		Top 20 pairs	
	No delay	1-day delay	No delay	1-day delay	No delay	1-day delay
Panel A: China Security Index 300 (CSI300) constituents						
$R_P\_EW\_CC$	17.05 (0)	0 (55.81)	16.28 (0)	0 (65.12)	18.60 (0)	0 (100)
$R_P\_EW\_FI$	15.50 (0)	0 (58.91)	14.73 (0)	0 (79.07)	25.58 (0)	0 (100)
$R_P\_VW\_CC$	13.18 (0)	0 (55.81)	12.40 (0)	0 (72.87)	14.73 (0)	0 (100)
$R_P\_VW\_FI$	7.75 (0)	0 (71.32)	11.63 (0)	0 (199.22)	26.36 (0)	0 (100)
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents						
$R_P\_EW\_CC$	0 (0)	0 (100)	0 (0)	0 (71.32)	0 (0)	0 (72.09)
$R_P\_EW\_FI$	0 (0)	0 (100)	0 (0)	0 (71.32)	0 (0)	0 (96.12)
$R_P\_VW\_CC$	0 (0)	0 (100)	0 (0)	0 (71.32)	0 (15.50)	0 (71.32)
$R_P\_VW\_FI$	0 (0)	0 (100)	0 (0)	0 (71.32)	0 (15.50)	0 (86.05)
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents						
$R_P\_EW\_CC$	0 (10.85)	0 (70.54)	0 (10)	0 (67.44)	9.30 (0)	0 (68.99)
$R_P\_EW\_FI$	0.78 (0)	0 (68.99)	23.26 (0)	0 (67.44)	27.13 (0)	0 (66.67)
$R_P\_VW\_CC$	0 (13.95)	0 (70.54)	0 (0)	0 (66.67)	12.40 (0)	0 (65.12)
$R_P\_VW\_FI$	0.78 (0)	0 (68.22)	23.26 (0)	0 (66.67)	27.13 (0)	0 (63.57)
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)						
$R_P\_EW\_CC$	41.09 (0)	0 (31.78)	45.74 (0)	0 (89.92)	21.71 (0)	0 (98.45)
$R_P\_EW\_FI$	42.64 (0)	0 (27.91)	37.21 (0)	0 (96.12)	21.71 (0)	0 (100)
$R_P\_VW\_CC$	41.86 (0)	0 (27.13)	55.04 (0)	0 (60.47)	24.03 (0)	0 (95.35)
$R_P\_VW\_FI$	44.19 (0)	0 (51.94)	50.39 (0)	0 (100)	22.48 (0)	0 (100)

Notes: This table presents the percentage of statistically significant profitable (unprofitable) subsamples from 01/1997 to 07/2017 based on a rolling-window regression for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The values in parentheses are the percentgage of statistically significant unprofitable subsamples. A statistically significant and positive (negative) intercept indicates the subsample is profitable (unprofitable). By setting the window length to 120-month, there are 129 subsamples. For the sake of brevity, we report the results at the 5% significance level only. $R_P\_EW\_CC$ is the Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is Value-Weighted Fully Invested scheme. All the values are in % term.

Table 6. Data-mining vs. genuine of the top 5 pairs

	No delay						1-day delay					
	Abnormal return			t-stat of abnormal return			Abnormal return			t-stat of abnormal return		
	Act α	Sim α	%	Act t_α	Sim t_α	%	Act α	Sim α	%	Act t_α	Sim t_α	%
Panel A: China Security Index 300 (CSI300) constituents												
1	-0.16	-0.62	0	-0.66	-2.70	0	-1.08	-0.67	99.6	-3.42	-2.48	97.13
2	-0.15	-0.53	0	-0.62	-2.30	0	-1.08	-0.56	100	-3.40	-2.12	99.96
3	-0.14	-0.48	0	-0.59	-2.08	0	-1.08	-0.51	100	-3.39	-1.92	100
4	-0.14	-0.44	0	-0.57	-1.92	0	-1.08	-0.47	100	-3.37	-1.77	100
5	-0.13	-0.42	0	-0.56	-1.80	0	-1.07	-0.44	100	-3.37	-1.66	100
10	-0.12	-0.32	0	-0.46	-1.40	0	-1.06	-0.33	100	-3.17	-1.28	100
20	-0.07	-0.22	0	-0.23	-0.94	0	-0.94	-0.21	100	-3.07	-0.83	100
30	-0.04	-0.14	0.01	-0.14	-0.62	0.01	-0.82	-0.13	100	-3.01	-0.52	100
40	0.03	-0.08	0	0.12	-0.36	0.01	-0.77	-0.07	100	-2.74	-0.27	100
50	0.06	-0.03	0.09	0.22	-0.12	0.19	-0.67	-0.01	100	-2.58	-0.04	100
60	0.08	0.02	1.42	0.30	0.10	3.73	-0.63	0.05	100	-2.45	0.18	100
70	0.16	0.07	0.04	0.59	0.31	0.66	-0.60	0.10	100	-2.22	0.40	100
80	0.28	0.12	0	1.02	0.53	0	-0.47	0.15	100	-1.72	0.62	100
90	0.41	0.17	0	1.61	0.76	0	-0.37	0.21	100	-1.49	0.86	100
100	0.50	0.23	0	1.86	1.02	0	-0.30	0.27	100	-1.14	1.12	100
110	0.54	0.30	0	1.98	1.34	0	-0.27	0.35	100	-1.01	1.45	100
120	0.57	0.39	0	2.21	1.80	1.60	-0.23	0.46	100	-0.93	1.94	100
125	0.57	0.48	3.24	2.27	2.21	37.63	-0.21	0.55	100	-0.87	2.37	100
126	0.58	0.51	8.64	2.28	2.33	56.17	-0.21	0.58	100	-0.86	2.51	100
127	0.60	0.54	16.46	2.39	2.49	62.26	-0.21	0.62	100	-0.86	2.68	100
128	0.60	0.58	39.52	2.40	2.71	81.68	-0.21	0.67	100	-0.86	2.92	100
129	0.61	0.67	71.70	2.41	3.12	95.63	-0.02	0.77	100	-0.78	3.36	100
Panel B: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)												
1	-0.13	-0.64	0	-0.69	-2.82	0	-0.77	-0.76	57.91	-3.52	-2.70	95.18
2	-0.11	-0.54	0	-0.59	-2.42	0	-0.73	-0.65	81.96	-3.41	-2.32	99.87
3	-0.10	-0.49	0	-0.56	-2.19	0	-0.73	-0.59	95.75	-3.36	-2.11	99.98
4	-0.10	-0.46	0	-0.53	-2.02	0	-0.73	-0.55	98.98	-3.33	-1.96	99.99
5	-0.10	-0.43	0	-0.52	-1.90	0	-0.73	-0.51	99.85	-3.33	-1.84	99.99
10	-0.09	-0.33	0	-0.46	-1.47	0	-0.69	-0.40	100	-3.16	-1.45	99.99
20	-0.03	-0.22	0	-0.16	-0.97	0	-0.61	-0.27	100	-2.87	-0.99	99.99
30	-0.01	-0.12	0	-0.05	-0.62	0	-0.55	-0.18	100	-2.42	-0.67	99.99
40	0.10	-0.07	0	0.43	-0.33	0	-0.51	-0.11	100	-2.0	-0.40	99.99
50	0.18	-0.01	0	0.81	-0.07	0	-0.46	-0.04	100	-1.83	-0.16	99.99
60	0.25	0.04	0	1.06	0.18	0	-0.44	0.02	100	-1.76	0.06	99.99
70	0.32	0.09	0	1.42	0.43	0	-0.42	0.07	100	-1.66	0.29	99.99
80	0.46	0.14	0	2.07	0.68	0	-0.39	0.13	100	-1.53	0.53	99.99
90	0.49	0.19	0	2.24	0.94	0	-0.38	0.19	100	-1.41	0.78	99.99
100	0.55	0.25	0	2.35	1.24	0	-0.35	0.25	100	-1.35	1.07	99.99
110	0.64	0.31	0	2.56	1.61	0	-0.33	0.33	100	-1.28	1.44	99.99
120	0.67	0.41	0	2.79	2.14	0.10	-0.30	0.44	100	-1.07	1.99	99.99
125	0.70	0.49	0	3.46	2.60	0.14	-0.24	0.53	100	-0.92	2.49	99.99
126	0.71	0.51	0.01	3.49	2.74	0.69	-0.22	0.55	100	-0.90	2.65	99.99
127	0.71	0.54	0.21	3.54	2.92	3.44	-0.22	0.59	100	-0.90	2.85	99.99
128	0.73	0.58	1.35	3.64	3.17	1.89	-0.20	0.63	100	-0.81	3.14	99.99
129	0.75	0.65	12.84	3.74	3.62	37.37	-0.17	0.71	100	-0.71	3.66	99.99

Notes: This table applies the modified Fama and French (2010) bootstrap simulation to the 129 subsamples to investigate whether the time-varying profitability of pairs trading is due to data-mining (luck) or real stock picking ability. The pairs trading portfolio return is based on the Equally-Weighted Committed Capital scheme. The column "Act α" is the actual abnormal return of pairs trading throughout the 129 subsamples by ascending order. The column "Sim α" is the average of the 10,000 simulated abnormal returns. The column "%" measures the percentage of simulated abnormal return above the actual abnormal return. The column "Act t_α" is the actual t-statistic of abnormal return throughout the 129 subsamples by ascending order. "Sim t_α" is the average of the 10,000 simulated t-statistics of abnormal returns. For the outperformance (underperformance) subsamples which is measured by actual α (or t_α), if simulated performance is greater than actual performance in less than 5% of the 10,000 simulations, at any given performance order, we reject the null hypothesis that the outperformance (underperformance) of pairs trading is due to good luck (poor stock picking skill) at 95% confidence level and infer that the strategy is genuine (bad luck). The simulated t_α are based on Newey-West standard error with 0 lag, as there is no autocorrelation in the bootstrapping process. Act α and Sim α are in % terms.

Table 7. Pairs trading performance in short-term market conditions

	CSI300 constituents		HSHKI constituents		HSAHP constituents		ALL	
	$Rm \geq 0$	$Rm < 0$	$Rm \geq 0$	$Rm < 0$	$Rm \geq 0$	$Rm < 0$	$Rm \geq 0$	$Rm < 0$
Top 5 pairs								
Abnormal	0.10	-0.59**	-0.51	-0.07	0.04	0.06	-0.24	0.33
Return	(0.33)	(0.29)	(0.32)	(0.42)	(0.33)	(0.31)	(0.34)	(0.25)
β_{R_M}	0.01	-0.08**	0.05	0.01	0.01	-0.02	0.07*	-0.04*
	(0.03)	(0.03)	(0.04)	(0.04)	(0.06)	(0.02)	(0.04)	(0.03)
Top 10 pairs								
Abnormal	-0.05	-0.58**	-0.27	1.49	-0.01	-0.01	-0.03	0.10
Return	(0.19)	(0.24)	(0.25)	(1.16)	(0.22)	(0.20)	(0.20)	(0.22)
β_{R_M}	0.03	-0.07***	0.01	0.05	0.04	-0.01	0.05	-0.04**
	(0.02)	(0.02)	(0.04)	(0.06)	(0.03)	(0.02)	(0.03)	(0.02)
Top 20 pairs								
Abnormal	-0.03	-0.33*	-0.29	1.11	-0.09	0.06	-0.12	-0.09
Return	(0.15)	(0.17)	(0.20)	(0.79)	(0.14)	(0.13)	(0.17)	(0.17)
β_{R_M}	0.03*	-0.06***	0.00	0.02	0.03*	-0.00	0.04*	-0.03**
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)

Notes: The pairs trading portfolio return is based on Equally-Weighted Committed Capital scheme under the scenario of no trading delay. Abnormal return is the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row "β_{R_M}" means the coefficient for the market factor. The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. Abnormal returns are in % terms.

Table 8. Pairs trading performance in longer-term market conditions

	CSI300 constituents		HSHKI constituents		HSAHP constituents		ALL	
	$LRm \geq 0$	$LRm < 0$	$LRm \geq 0$	$LRm < 0$	$LRm \geq 0$	$LRm < 0$	$LRm \geq 0$	$LRm < 0$
Top 5 pairs								
Abnormal	-0.06	0.10	0.01	-0.23	0.03	0.16	0.12	0.52**
return	(0.31)	(0.19)	(0.30)	(0.22)	(0.32)	(0.19)	(0.21)	(0.22)
β_{R_M}	0.00	0.01	0.01	0.02	-0.00	-0.01	-0.01	-0.03
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
Top 10 pairs								
Abnormal	-0.08	0.05	1.53	-0.09	0.10	0.20	0.12	0.29**
return	(0.19)	(0.16)	(1.16)	(0.18)	(0.17)	(0.15)	(0.17)	(0.15)
β_{R_M}	0.02	-0.02	-0.08	-0.02	0.02	-0.01	0.02	-0.03*
	(0.02)	(0.02)	(0.07)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Top 20 pairs								
Abnormal	-0.03	0.13	0.81	0.00	0.09	0.14	-0.02	0.18
return	(0.17)	(0.10)	(0.77)	(0.14)	(0.10)	(0.11)	(0.17)	(0.13)
β_{R_M}	0.01	-0.02	-0.06	-0.03*	0.01	-0.01	0.01	-0.02
	(0.02)	(0.02)	(0.05)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)

Notes: The pairs trading portfolio return is based on Equally-Weighted Committed Capital scheme under the scenario of no trading delay. LRm is defined as the moving average of previous 12-month market excess return. Abnormal return is the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row "β_{R_M}" means the coefficient for the market factor. The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. Abnormal returns are in % terms.

Table 9. Profitability of pairs trading as return enhancer

	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$
Panel A: Top 5 pairs				
CSI300 constituents	0.01	0.04	0.01	0.03
	(0.10)	(0.16)	(0.09)	(0.16)
HSHKI constituents	-0.05	-0.16	-0.06	-0.17
	(0.09)	(0.18)	(0.08)	(0.18)
HSAHP constituents	0.05	0.25	0.06	0.26*
	(0.09)	(0.15)	(0.09)	(0.15)
ALL	0.20**	0.36**	0.21***	0.37***
	(0.08)	(0.14)	(0.08)	(0.14)
Panel B: Top 10 pairs				
CSI300 constituents	0.01	0.06	0.01	0.04
	(0.07)	(0.11)	(0.06)	(0.10)
HSHKI constituents	0.30	0.31	0.70	0.66
	(0.29)	(0.36)	(0.61)	(0.64)
HSAHP constituents	0.07	0.34**	0.07	0.34**
	(0.05)	(0.14)	(0.05)	(0.14)
ALL	0.14**	0.28**	0.16***	0.30***
	(0.06)	(0.11)	(0.06)	(0.11)
Panel C: Top 20 pairs				
CSI300 constituents	0.04	0.12	0.05	0.12
	(0.05)	(0.09)	(0.05)	(0.09)
HSHKI constituents	0.17	-0.01	0.57	0.45
	(0.19)	(0.04)	(0.53)	(0.58)
HSAHP constituents	0.05	0.36***	0.05	0.36***
	(0.04)	(0.14)	(0.03)	(0.17)
ALL	0.06	0.17	0.07	0.17*
	(0.05)	(0.11)	(0.05)	(0.10)

Notes: This table presents the abnormal return of an index enhancer strategy, combination of pairs trading portfolio return under the scenario of no trading delay with the broad market return, by using the Treynor and Black (1973) model. $R_P\_EW\_CC$ is the Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is Value-Weighted Fully Invested scheme. Abnormal return is the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. Abnomral returns are in % term.

Figure 1. Cumulative monthly return indices (in % term) for a portfolio of top 5 pairs

HSAHP

ALL
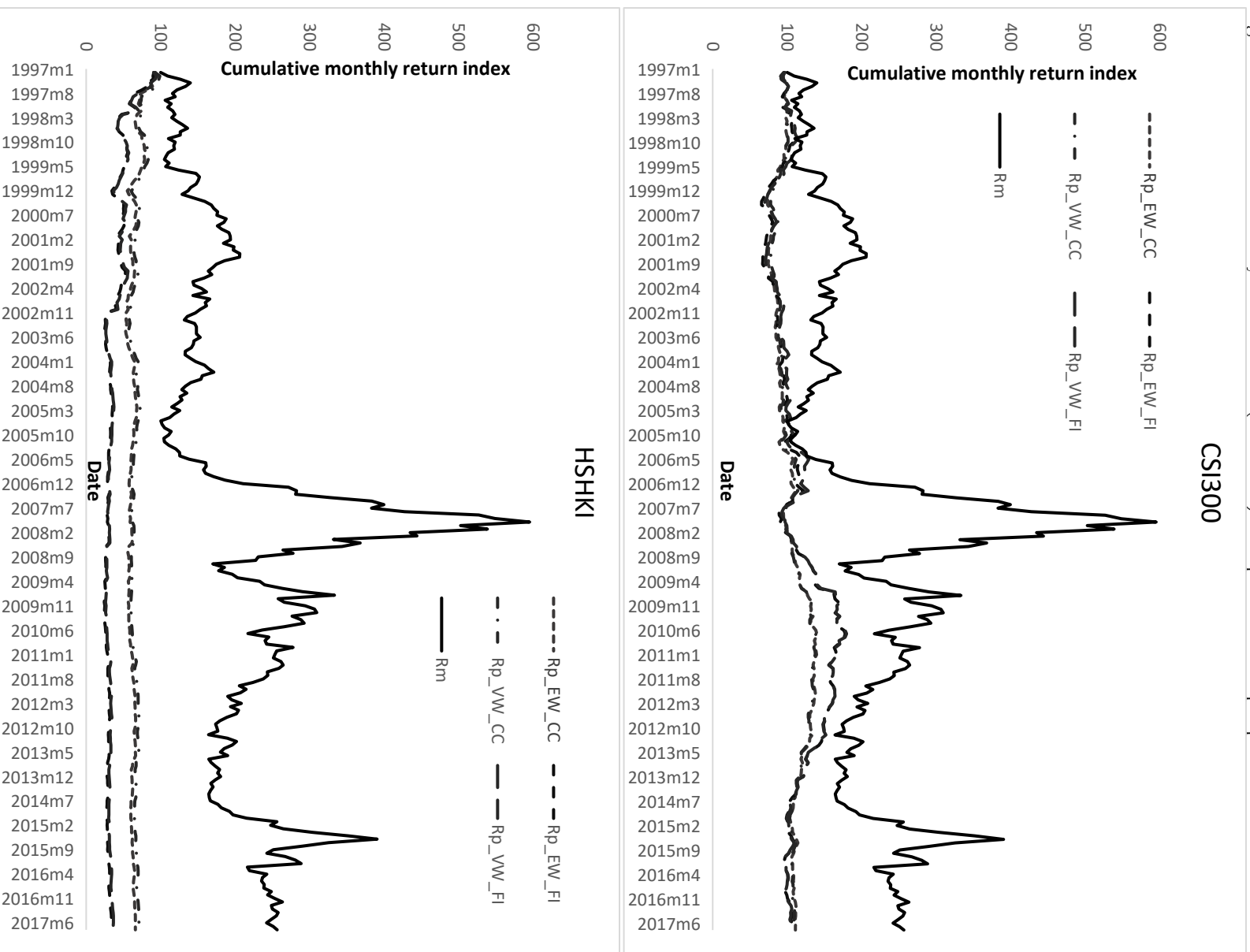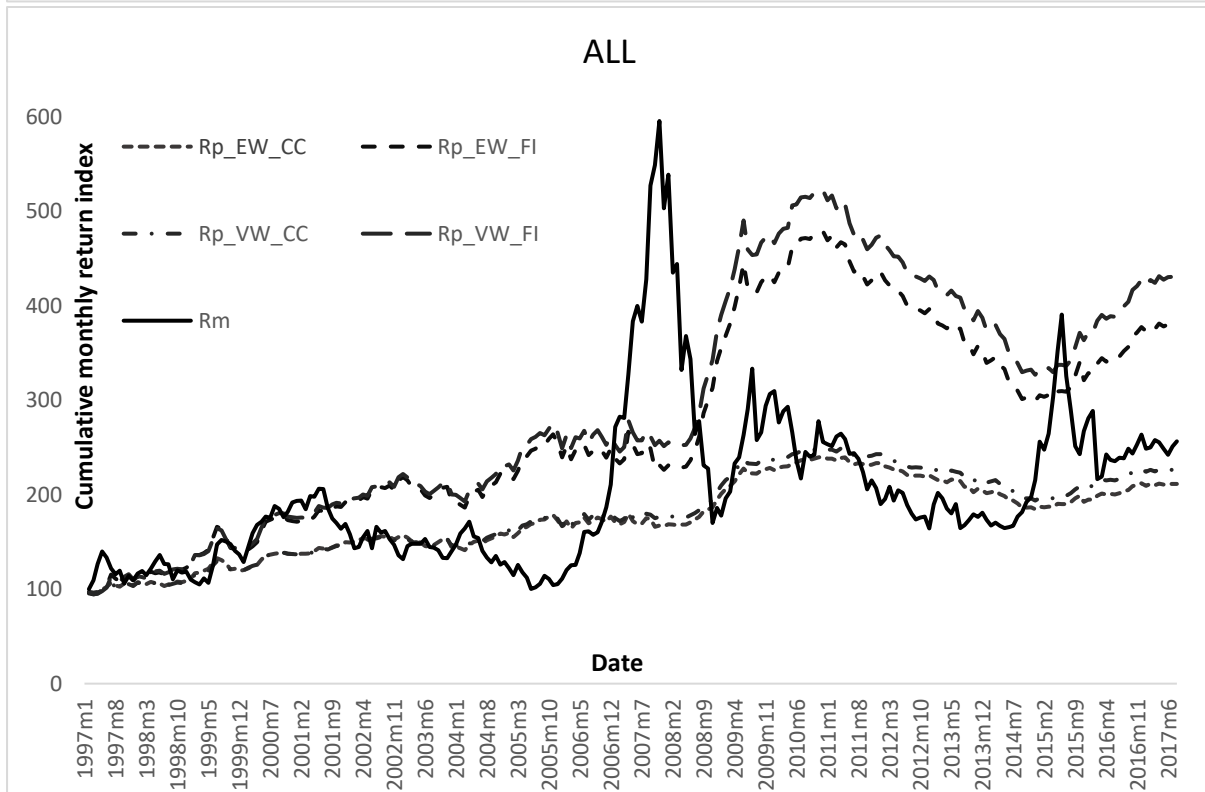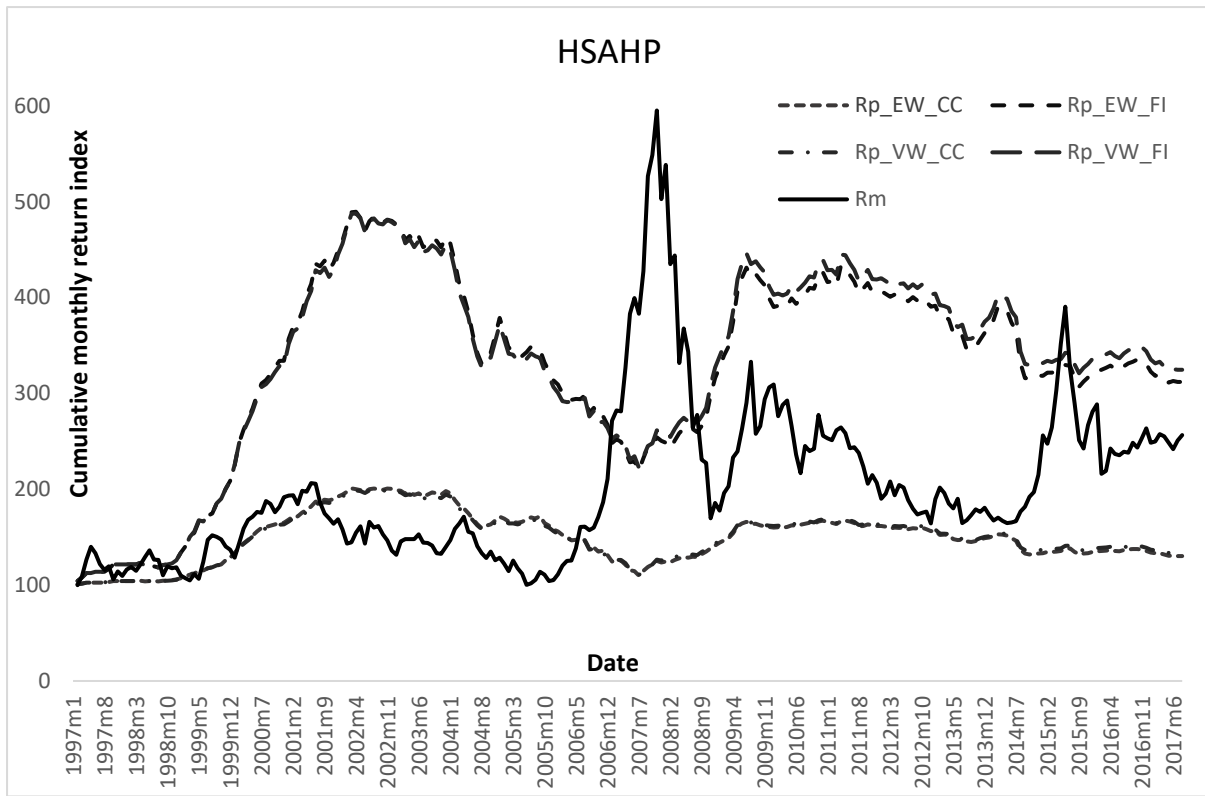
Notes: CSI300 stands for the China Security Index 300 constituents. HSHKI stands for the Hang Seng Stock Connect Hong Kong Index large- and mid-cap constituents. HSAHP stands for the Hang Seng Stock Connect China AH Premium Index constituents. ALL stands for the sum of CSI300 constituents, HSHKI large- and mid-cap constituents, and HSAHP constituents. $R_P\_EW\_CC$ stands for the pairs trading portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ stands for the pairs trading portfolio return based on Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ stands for the pairs trading portfolio return based on Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ stands for the pairs trading portfolio return based on Value-Weighted Fully Invested scheme. Rm is the excess return of the market over the risk-free rate. All the returns are based on the trading scenario of no trading delay. The returns are in % term.

Figure 2. The time-varying *t*-statistic of abnormal return for a portfolio of top 5 pairs
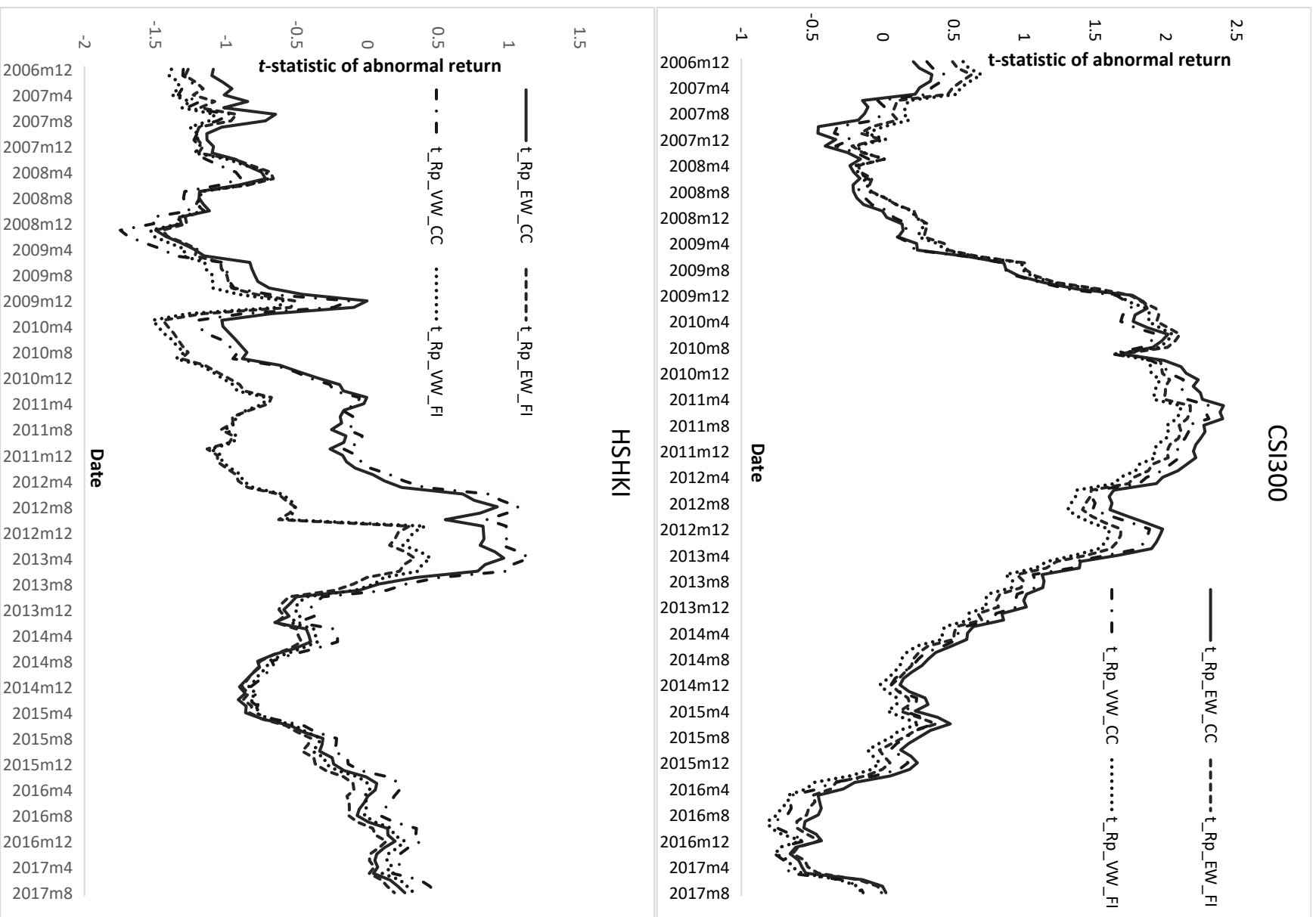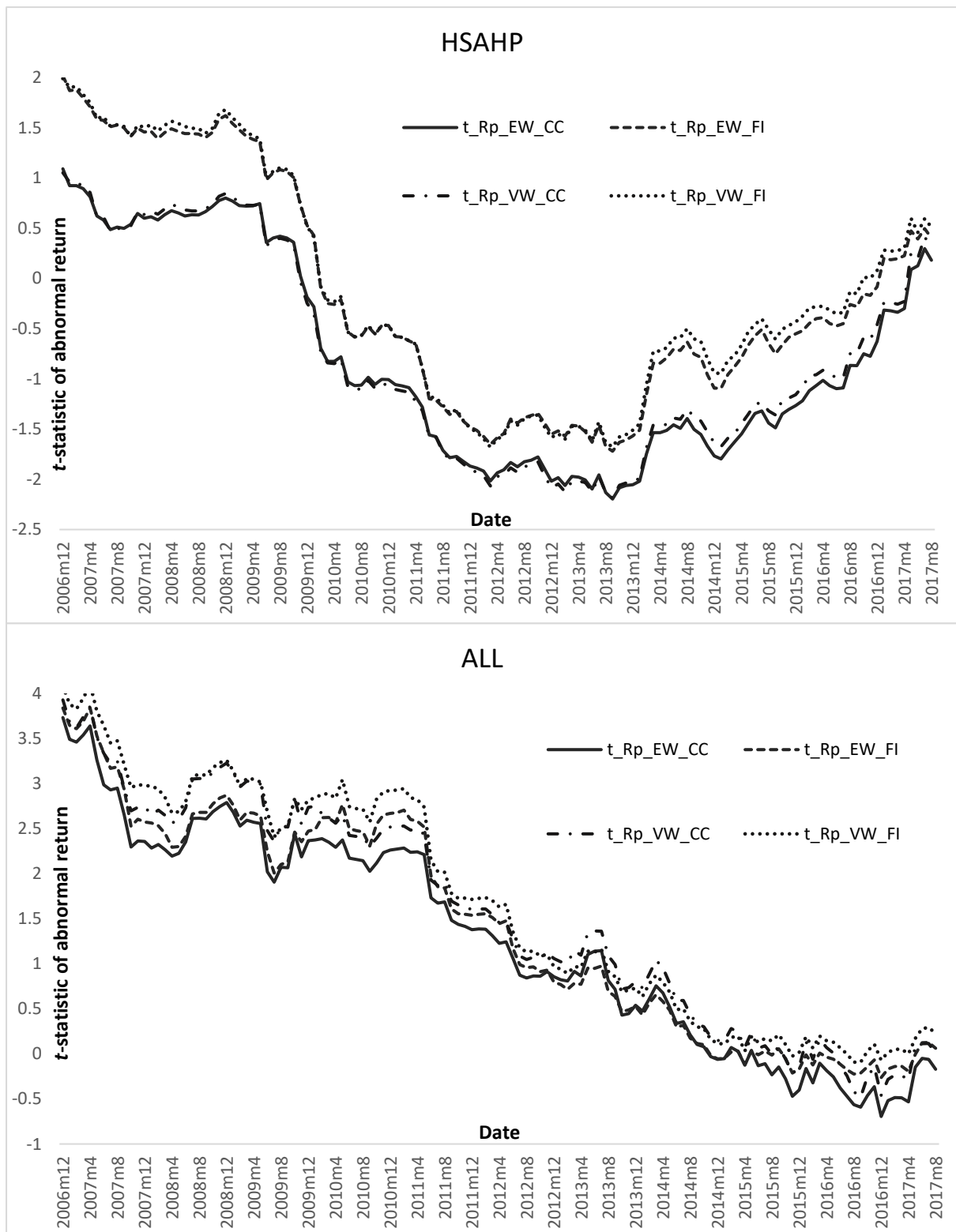
Notes: This figure presents the *t*-statistic for the regression intercept for the monthly return series (under the scenario of no trading delay) against Fama-French 5-factor plus momentum and short-term reversal factors with window length 120-month and rolling forward 1-month each time. $R_P\_EW\_CC$ stands for the pairs trading portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ stands for the pairs trading portfolio return based on Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ stands for the pairs trading portfolio return based on Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ stands for the pairs trading portfolio return based on Value-Weighted Fully Invested scheme. The critical values for the two-tail *t*-test are ±1.645, ±1.96, and ±2.576 at the 10%, 5% and 1% significance level, respectively. A statistically significant and positive *t*-statistic for the regression intercept would reject the hypothesis that the market is efficient at any given subsample.

References

Auer, B. R. and Schuhmacher, F. (2015) 'Liquid betting against beta in Dow Jones Industrial Average stocks', *Financial Analysts Journal*, 71(6), pp. 30-43.

Bae, K.-H., Bailey, W. and Mao, C. X. (2006) 'Stock market liberalization and the information environment', *Journal of International Money and Finance*, 25(3), pp. 404-428.

Baele, L. (2005) 'Volatility spillover effects in European equity markets', *Journal of Financial and Quantitative Analysis*, 40(2), pp. 373-401.

Baker, M., Bradley, B. and Taliaferro, R. (2014) 'The Low-risk anomaly: A decomposition into micro and macro effects', *Financial Analysts Journal*, 70(2), pp. 43-58.

Blake, D., Caulfield, T., Ioannidis, C. and Tonks, I. (2017) 'New evidence on mutual fund performance: A comparison of alternative bootstrap methods', *Journal of Financial and Quantitative Analysis*, 52(3), pp. 1279-1299.

Bowen, D. A. and Hutchinson, M. C. (2016) 'Pairs trading in the UK equity market: Risk and return', *European Journal of Finance*, 22(14), pp. 1363-1387.

Bowen, D. A., Hutchinson, M. C. and O'Sullivan, N. (2010) 'High-frequency equity pairs trading: Transaction costs, speed of execution, and patterns in returns', *Journal of Trading*, 5(3), pp. 31-38.

Broussard, J. P. and Vaihekoski, M. (2012) 'Profitability of pairs trading strategy in an illiquid market with multiple share classes', *Journal of International Financial Markets, Institutions and Money*, 22(5), pp. 1188-1201.

Campbell, J. Y., Lo, A. W. C. and MacKinlay, A. C. (1997) *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

Cuthbertson, K., Nitzsche, D. and O'Sullivan, N. (2008) 'UK mutual fund performance: Skill or luck?', *Journal of Empirical Finance*, 15(4), pp. 613-634.

Do, B. and Faff, R. (2010) 'Does simple pairs trading still work?', *Financial Analysts Journal*, 66(4), pp. 83-95.

Do, B. and Faff, R. (2012) 'Are pairs trading profits robust to trading costs', *Journal of Financial Research*, 35(2), pp. 261-287.

Elliott, R. J., Van Der Hoek , J. and Malcolm, W. P. (2005) 'Pairs trading', *Quantitative Finance*, 5(3), pp. 271-276.

Fama, E. F. and French, K. R. (2010) 'Luck versus skill in the cross-section of mutual fund returns', *Journal of Finance*, 65(5), pp. 1915-1947.

Fama, E. F. and French, K. R. (2015) 'A five-factor asset pricing model', *Journal of Financial Economics*, 116(1), pp. 1-22.

Financial Times (2013) *End of the road for China's 'B' market* Available at: https://www.ft.com/content/254b3b6e-5a2a-11e2-a02e-00144feab49a (Accessed: 9th January, 2013).

Gallefoss, K., Hansen, H. H., Haukaas, E. S. and Molnár, P. (2015) 'What daily data can tell us about mutual funds: Evidence from Norway', *Journal of Banking and Finance*, 55, pp. 117-129.

Gatev, E., Goetzmann, W. N. and Rouwenhorst, K. G. (2006) 'Pairs trading: Performance of a relative-value arbitrage rule', *Review of Financial Studies*, 19(3), pp. 797-827.

Grossman, S. J. and Stiglitz, J. E. (1980) 'On the impossibility of informationally efficient markets', *American Economic Review*, 70(3), pp. 393-408.

Hakkio, C. S. and Rush, M. (1991) 'Is the budget deficit "too large"?', *Economic Inquiry*, 29(3), pp. 429-445.

Hiremath, G. S. and Narayan, S. (2016) 'Testing the adaptive market hypothesis and its determinants for the Indian stock markets', *Finance Research Letters*, 19(Supplement C), pp. 173-180.

Jacobs, H. and Weber, M. (2015) 'On the determinants of pairs trading profitability', *Journal of Financial Markets*, 23, pp. 75-97.

Kim, J. H., Shamsuddin, A. and Lim, K.-P. (2011) 'Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data', *Journal of Empirical Finance*, 18(5), pp. 868-879.

Kosowski, R., Timmermann, A., Wermers, R. and White, H. A. L. (2006) 'Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis', *Journal of Finance*, 61(6), pp. 2551-2595.

Krauss, C. (2017) 'Statistical arbitrage pairs trading strategies: Review and outlook', *Journal of Economic Surveys*, 31(2), pp. 513-545.

Li, M. L., Chui, C. M. and Li, C. Q. (2014) 'Is pairs trading profitable on China AH-share markets?', *Applied Economics Letters*, 21(16), pp. 1116-1121.

Lo, A. W. (2004) 'The Adaptive Markets Hypothesis', *Journal of Portfolio Management*, 30(5), pp. 15-29.

Lo, A. W. (2005) 'Reconciling efficient markets with behavioral finance: The Adaptive Markets Hypothesis', *Journal of Investment Consulting*, 7(2), pp. 21-44.

Lo, A. W. and MacKinlay, A. C. (1990) 'An econometric analysis of nonsynchronous trading', *Journal of Econometrics*, 45(1), pp. 181-211.

Marshall, B. R., Nguyen, N. H. and Visaltanachoti, N. (2013) 'ETF arbitrage: Intraday evidence', *Journal of Banking and Finance*, 37(9), pp. 3486-3498.

Meyer, S., Schmoltzi, D., Stammschulte, C., Kaesler, S., Loos, B. and A., H. (2012) 'Just unlucky? – A bootstrapping simulation to measure skill in individual investors' investment performance', *Working Paper Goethe University Frankfurt*.

Newey, W. K. and West, K. D. (1987) 'A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica*, 55(3), pp. 703-708.

Rad, H., Low, R. K. Y. and Faff, R. (2016) 'The profitability of pairs trading strategies: Distance, cointegration and copula methods', *Quantitative Finance*, 16(10), pp. 1541-1558.

Schotman, P. C. and Zalewska, A. (2006) 'Non-synchronous trading and testing for market integration in Central European emerging markets', *Journal of Empirical Finance*, 13(4), pp. 462-494.

Stiglitz, J. E. (2003) 'Globalization and growth in emerging markets and the New Economy', *Journal of Policy Modeling*, 25(5), pp. 505-524.

Sun, Q., Tong, W. H. S. and Yan, Y. (2009) 'Market liberalization within a country', *Journal of Empirical Finance*, 16(1), pp. 18-41.

Timmermann, A. and Granger, C. W. J. (2004) 'Efficient market hypothesis and forecasting', *International Journal of Forecasting*, 20(1), pp. 15-27.

Tourin, A. and Yan, R. (2013) 'Dynamic pairs trading using the stochastic control approach', *Journal of Economic Dynamics and Control*, 37(10), pp. 1972-1981.

Treynor, J. L. and Black, F. (1973) 'How to use security analysis to improve portfolio selection', *Journal of Business*, 46(1), pp. 66-86.

Urquhart, A. and Hudson, R. (2013) 'Efficient or adaptive markets? Evidence from major stock markets using very long run historic data', *International Review of Financial Analysis*, 28, pp. 130-142.

Urquhart, A. and McGroarty, F. (2014) 'Calendar effects, market conditions and the Adaptive Market Hypothesis: Evidence from long-run U.S. data', *International Review of Financial Analysis*, 35, pp. 154-166.

Urquhart, A. and McGroarty, F. (2016) 'Are stock markets really efficient? Evidence of the adaptive market hypothesis', *International Review of Financial Analysis*, 47, pp. 39-49.

Vidyamurthy, G. (2004) *Pairs Trading: Quantitative Methods and Analysis*. New Jersey: Wiley.

White, H. (2000) 'A reality check for data snooping', *Econometrica*, 68(5), pp. 1097-1126.

Xie, W., Liew, R. Q., Wu, Y. and Zou, X. (2016) 'Pairs trading with copulas', *Journal of Trading*, 11(3), pp. 41-52.

Appendices

Table A1. Summary statistics and correlations for monthly factor returns

Panel A: Means, standard deviations, and *t*-statistics of factor returns							
	R_M	SMB	HML	RMW	CMA	MOM	ST_REV
Mean	0.76	0.99	0.34	0.22	-0.11	-1.80	0.85
Std. Dev.	8.35	4.23	3.87	4.06	3.35	5.70	4.04
t-statistic	1.41	3.64	1.36	0.83	-0.51	-4.90	3.25

Panel B: Correlations between different factors							
	R_M	SMB	HML	RMW	CMA	MOM	ST_REV
R_M	1.00						
SMB	0.01	1.00					
HML	0.15**	-0.26***	1.00				
RMW	-0.20***	-0.36***	-0.17***	1.00			
CMA	-0.01	0.21***	0.20***	-0.67***	1.00		
MOM	-0.11*	-0.50***	-0.11	0.43***	-0.32***	1.00	
ST_REV	-0.07	-0.35***	0.16**	0.12*	-0.09	-0.01	1.00

Notes: Mean and Std. Dev. are the mean and standard deviation of the return, and *t*-statistic is ratio of the mean return over its standard error. R_M denotes value-weighted monthly excess returns, including all A-share stocks listed on Shanghai and Shenzhen stock exchanges. We use the 3-month China time deposit rate to proxy the risk-free rate. At the end of June each year, stocks are sorted into two size groups using the main-board market median market capitalization as the breakpoint. At the same time, stocks are also sorted independently into three book-to-market equity, operating profitability, and investment groups, using the main-board market 30th and 70th percentiles of book-to-market, operating profitability, and investment. The Fama-French five factors (2x3) are constructed using the six value-weight portfolios formed on size and book-to-market, the six value-weight portfolios formed on size and operating profitability, and the six value-weight portfolios formed on size and investment. SMB (Small Minus Big) is the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios. HML (High Minus Low) is the average return on the two value portfolios minus the average return on the two growth portfolios. RMW (Robust Minus Weak) is the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios. CMA (Conservative Minus Aggressive) is the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios. Similarly, we use six value-weight portfolios formed on size and prior (2-12) returns to construct MOM (Momentum). The momentum portfolios, which are formed monthly, are the intersections of two portfolios formed on size (market equity) and three portfolios formed on prior (2-12) return. The monthly size breakpoint is the median main-board market capitalization. The monthly prior (2-12) return breakpoints are the 30th and 70th main-board market percentiles. Mom is the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios. To be included in a portfolio for month t (formed at the end of month t-1), a stock must have a price for the end of month t-13 and a good return for t-2. We use six value-weight portfolios formed on size and prior (1-1) returns to construct ST_REV (short-term reversal). The short-term reversal portfolios, which are formed monthly, are the intersections of two portfolios formed on size (market equity, ME) and three portfolios formed on prior (1-1) return. The monthly prior (1-1) return breakpoints are the 30th and 70th main-board market percentiles. ST_Rev is the average return on the two low prior return portfolios minus the average return on the two high prior return portfolios. The portfolios used to construct factors include all A-share stocks listed on Shanghai and Shenzhen stock exchanges. We exclude stocks with negative or missing book values. See Ken French's webpage (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) for more detailed discussion on factor construction. ***, ** and * denotes statistically significant at the 1%, 5% and 10%, respectively.

Table A2. Key dates in the mainland China and Hong Kong stock markets

Date	Description
1891	The Association of Stockbrokers in Hong Kong was established in 1891. It was renamed to the Hong Kong Stock Exchange in 1914.
19/12/1990	Shanghai Stock Exchange (SHSE) was established on 26/11/1990 and was in operation on 19/12/1990.
16/04/1991	Shanzhen Stock Exchange (SZSE) starts trail operation on 01/12/1990 and was officially approved on 16/04/1991.
19/02/2001	B-shares, quoted in US Dollars on SHSE and Hong Kong Dollars on SZSE, are exclusively available to foreign investors until 19/02/2001; thereafter, also available to mainland China investors.
09/07/2003	A-shares are stocks, quoted in Chinese Yuan, exclusively available to mainland China citizens and home institutions until 09/07/2003; thereafter, also available to non-resident investors under the Qualified Foreign Institutional Investor (QFII) scheme.
08/04/2005	China Securities Index Company Ltd launches the China Securities Index (CSI 300).
09/07/2007	Hang Seng Indexes Company Ltd launches the Hang Seng Stock Connect China AH Premium Index (HSAHP).
31/03/2010	Short-selling was officially permitted in A-shares market.
17/11/2014	Shanghai and Hong Kong Stock exchanges launch the Shanghai-Hong Kong Stock Connect scheme.
05/12/2016	Shenzhen and Hong Kong Stock exchanges launch the Shenzhen-Hong Kong Stock Connect scheme. Hang Seng Indexes Company Ltd launches the Hang Seng Stock Connect Hong Kong Index (HSHKI).

Table A3. Monthly net excess returns of pairs trading for a portfolio of all pairs

	No delay				1-day delay			
	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$	$R_P\_EW\_CC$	$R_P\_EW\_FI$	$R_P\_VW\_CC$	$R_P\_VW\_FI$
Panel A: China Security Index 300 (CSI300) constituents								
Average excess return	-0.00	0.08	0.00	0.08	-0.01	-1.10***	-0.00***	-0.90***
	(0.00)	(0.33)	(0.00)	(0.33)	(0.00)	(0.40)	(0.00)	(0.37)
Skewness	2.11	2.16	2.31	2.31	-0.83	2.24	0.26	2.63
Kurtosis	10.37	9.97	11.29	10.95	7.76	11.72	4.74	14.08
Jarque-Bera statistic	123.2***	115.3***	153.9***	144.6***	43.36***	164.3***	5.63*	257.0***
Minimum	-0.00	-4.07	-0.00	-4.03	-0.01	-5.81	-0.00	-5.39
Maximum	0.01	9.12	0.01	9.18	0.46	10.41	0.01	10.56
Abnormal return	0.00	0.46	0.00	0.29	-0.00***	-0.89**	-0.00	-0.87***
	(0.00)	(0.62)	(0.00)	(0.74)	(0.00)	(0.42)	(0.13)	(0.18)
β_{R_M}	0.00	0.07	0.00	-0.06	0.01	0.09*	0.01	0.08**
	(0.00)	(0.04)	(0.00)	(0.04)	(0.02)	(0.04)	(0.02)	(0.04)
Panel B: Hang Seng Stock Connect Hong Kong Index (HSHKI) large- and mid-cap constituents								
Average excess return	0.00	0.05	0.00	0.59	-0.00***	-1.65***	0.00***	-1.73***
	(0.00)	(0.23)	(0.01)	(0.49)	(0.00)	(0.31)	(0.00)	(0.42)
Skewness	1.98	1.01	14.04	7.41	0.05	0.78	8.83	3.97
Kurtosis	12.19	14.16	207.8	76.75	5.58	11.75	88.98	49.12
Jarque-Bera statistic	948.8***	1,264***	4,200***	5,600***	43.72***	516.6***	5,100***	1400***
Minimum	-0.01	-14.48	-0.01	-17.13	-0.01	-16.14	-0.01	-23.91
Maximum	0.01	22.13	0.46	85.74	0.01	20.88	0.32	46.10
Abnormal return	0.00	-0.06	0.00	0.50	-0.00***	-1.49***	-0.00***	-1.83***
	(0.00)	(0.39)	(0.00)	(0.81)	(0.00)	(0.54)	(0.00)	(0.57)
β_{R_M}	0.01	0.02	-0.01	0.04	-0.00	-0.03	0.00	0.07
	(0.01)	(0.03)	(0.02)	(0.05)	(0.00)	(0.04)	(0.00)	(0.06)
Panel C: Hang Seng Stock Connect China AH Premium Index (HSAHP) constituents								
Average excess return	0.00	0.78***	0.00	0.77***	-0.00***	-0.36**	-0.00**	-0.27
	(0.00)	(0.17)	(0.00)	(0.17)	(0.00)	(0.17)	(0.00)	(0.21)
Skewness	0.71	0.88	0.75	0.91	-0.81	0.29	5.76	3.51
Kurtosis	6.30	4.76	6.47	4.72	9.52	4.67	69.30	33.61
Jarque-Bera statistic	127.0***	60.6***	140.5***	61.48***	443.9***	30.7***	4,500***	9,699***
Minimum	-0.01	-5.92	-0.01	-5.56	-0.03	-8.03	-0.02	-7.12
Maximum	0.02	11.21	0.02	11.47	0.02	11.16	0.07	29.20
Abnormal return	0.00*	0.81***	0.00*	0.81***	-0.00***	-0.27	-0.00**	-0.11
	(0.00)	(0.26)	(0.00)	(0.26)	(0.00)	(0.23)	(0.00)	(0.28)
β_{R_M}	0.00	0.02	0.00	0.02	-0.00	-0.01	0.00	0.02
	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.03)	(0.00)	(0.03)
Panel D: ALL (CSI300 constituents + HSHKI large- and mid-cap constituents + HSAHP constituents)								
Average excess return	0.00	0.18**	0.01***	0.19*	-0.00***	-1.12***	-0.00***	-0.95***
	(0.00)	(0.06)	(0.00)	(0.31)	(0.00)	(0.32)	(0.00)	(0.30)
Skewness	0.40	1.73	0.39	1.64	-1.34	1.02	-0.84	1.34
Kurtosis	6.80	8.10	6.51	8.16	6.53	6.15	3.94	8.23
Jarque-Bera statistic	25.05**	63.25***	21.52***	62.37***	33.6***	24.03***	6.34**	59.06***
Minimum	-0.01	-4.01	-0.00	-4.13	-0.02	-5.78	-0.01	-5.73
Maximum	0.01	7.53	0.01	7.28	0.01	6.20	0.01	6.47
Abnormal return	0.02*	0.50*	0.00*	0.52	-0.00**	-1.96***	-0.00**	-1.82***
	(0.01)	(0.24)	(0.00)	(0.49)	(0.00)	(0.56)	(0.00)	(0.53)
β_{R_M}	-0.00	0.00	-0.00	0.02	-0.00**	-0.02	-0.00**	-0.00
	(0.00)	(0.03)	(0.01)	(0.03)	(0.00)	(0.04)	(0.00)	(0.04)

Notes: This table displays results of compounded monthly return series from 01/1997 to 07/2017. $R_P\_EW\_CC$ is the portfolio return based on Equally-Weighted Committed Capital scheme. $R_P\_EW\_FI$ is the Equally-Weighted Fully Invested scheme. $R_P\_VW\_CC$ is the Value-Weighted Committed Capital scheme. $R_P\_VW\_FI$ is the Value-Weighted Fully Invested scheme. The null hypothesis for the Jarque-Bera test is that the variable is normally distributed. The row labelled "Abnormal return" presents the estimated regression intercept for the monthly return series against Fama-French 5-factor plus momentum and short-term reversal factors. The row labelled "β_{R_M}" is the coefficient for the market factor. The authors construct the Chinese-version of risk factors by following Gatev *et al.* (2006) and Fama and French (2015). The values in parentheses are the Newey-West standard errors. *** significant at 1% level. ** significant at 5% level. * significant at 10% level. The returns are in % term.