Stock Market Openness and Market Quality: Evidence from the Shanghai-Hong Kong Stock Connect Program*

Li Xing
University of Victoria

Ke Xu†
University of Victoria

Xinwei Zheng
Deakin University

Xuekui Zhang
University of Victoria

April 29, 2019

Abstract

This paper studies the impact of capital market openness on high frequency market quality in China. The Shanghai-Hong Kong Stock Connect (SHHKConnect) program opens China's stock market to foreign investors and offers a natural experiment to investigate this question. Using a difference-in-differences approach, we find that market liberalization leads to lower bid-ask spreads, higher market depth, higher short-term volatility and higher effective spreads. Our findings imply that opening the markets to more sophisticated foreign investors is associated with higher competition and more cross-market arbitrage activities, that improve displayed liquidity, but increase effective spreads and short-term volatility of connected stocks.

Keywords: capital market openness, competition, cross-market arbitrage, bid-ask spread, effective spread, short-term volatility, Shanghai-Hong Kong Stock Connect

*This work is supported by the Social Sciences and Humanities Research Council (SSHRC) Grants 430-2018-00557, the Natural Sciences and Engineering Research Council (NSERC) Discovery Grants RGPIN-2017-04722, the Canada Research Chair 950-231363, and the NSERC Post Doctoral Fellowship PDF-471667-2015. First version: September 2018.

†Corresponding author. E-mail: kexu@uvic.ca

JEL Classification: G10.

1 Introduction

The liberalization of financial markets through the opening up of domestic stock markets to foreign investors has controversial effects on different aspects of the economy. Academic studies find that market liberalization leads to lower costs of capital by allowing risk sharing (Henry (2000) and Bekaert and Harvey (2000)), an improved information environment (Bae et al. (2006)), better market quality (Sun et al. (2009)), faster productivity growth (Bekaert et al. (2011) and Larrain and Stumpner (2017)), reduced agency problems and enhanced governance quality (Doidge et al. (2004)). However, market liberalization may also have some unintended consequences. Stiglitz (2010) argues that the possible contagion effect of disturbances spilling over from developed markets destabilizes the capital markets in emerging economies. Ng (2000) and Baele (2005) find significant volatility spillovers from Japan and the US to six Pacific-Basin equity markets and from the US market to European equity markets.

In this paper, we provide new evidence of capital market liberalization by examining its effect on high frequency market quality using the Shanghai-Hong Kong Stock Connect program (SHHKConnect). Recently, the Chinese government has made a sequence of policies to open up its capital markets to foreign investors, creating an ideal environment for examining the impact of increased foreign portfolio investment in developing equity markets. Our main focus is the influence of increased foreign investment activities on four market quality variables: bid-ask spread, effective spread, market depth and short-term volatility. Bid-ask spread and market depth are measures of displayed liquidity in the limit order market. Effective spreads reflect investors' actual trading costs. Short-term volatility is a good indicator for high frequency arbitrage activities.

Capital market openness may affect market quality in two ways. First, opening the market to more foreign investors increases competition for liquidity provision (competition channel). As competition increases, investors undercut each other's limit orders more aggressively to obtain a better chance for a trade to take place. Theoretical models of liquidity provision argue that the bid-ask spread decreases and market depth increases as the level of

competition increases (Ho and Stoll (1983), Dutta and Madhavan (1997) and Brogaard and Garriott (2014)). Second, with improved connectivity, more arbitrage opportunities become available. The more sophisticated foreign investors will take advantage of these arbitrage opportunities by trading fast on both the Shanghai Stock Exchange (SSE) and Hong Kong Stock Exchange. These arbitrage opportunities impose a cost to liquidity provision because stale quotes can be picked off by faster high frequency arbitragers. Theory predicts that more cross market arbitrage activities increase adverse selection (adverse selection channel), therefore putting upward pressure on bid-ask spreads, short-term volatility, and effective spreads (Glosten and Milgrom (1985), Budish et al. (2015), Foucault et al. (2017), Biais et al. (2015), Xu (2018), Hasbrouck (2018) and Zhang (2010)). Hence, market liberalization has an ambiguous effect on bid-ask spreads, depending on which mechanism dominates. Our paper aims to empirically investigate this theoretical ambiguity and disentangle which channel dominates in China's stock market. In particular, we are interested in the following research questions: (i) whether the displayed liquidity in the limit order book is improved, (ii) how investors' trading costs change, and (iii) whether we observe evidence of more high frequency arbitrage activities.

As China accelerated the opening of its capital market, more international investors, especially quantitative investors, or "quants", became more interested in investing in China. Considering that it becomes more difficult to make money in ferociously competitive and efficient developed markets, like the US, China's retail-dominated stock market could become "the industry's new Klondike with market gold readily available for mining" (Financial Times August 2, 2018).[1] Domestic investors in China are well aware of this situation and are concerned that opening the market to foreign investors will increase cross-market arbitrage activities, which intensifies adverse selection. In this paper, we address these concerns by studying the net effect of competition and adverse selection channels on the trading costs of investors.

On November 17, 2014, the Chinese government initiated SHHKConnect program, which allows investors in mainland China and Hong Kong to trade and settle eligible stocks listed on

[1]China's stock market has long been dominated by retail investors, who account for roughly 70% of all trading in China.

the other market via the exchange and clearing house in their home markets. SHHKConnect provides an ideal setting to investigate the effect of capital market openness on stock market quality, which may be of interest to both domestic and international investors, as well as policy makers in China. To investigate these questions, we utilize the high frequency order-level data on all firms listed on the SSE in the China Stock Market and Accounting Research (CSMAR) database, which provides real-time information about orders and executions on the SSE with millisecond timestamps. We employ identification tests using the difference-in-differences approach. We study the exogenous variation in market quality generated by SHHKConnect and show that stocks that are eligible for the connect program experience significant different changes in market quality compared to stocks that are not eligible for the connect program.

Following Hasbrouck and Saar (2013), we use three measures of market liquidity, the bid-ask spread, the effective spread and market depth, plus a measure of short-term volatility, to represent different aspects of market quality. The first measure is the time weighted average quoted spread (best ask price minus best bid price) on the SSE in a 10-minute interval. The second measure is the value-weighted average effective spread (or total price impact) of all trades on the SSE during a 10-minute interval, where the effective spread is defined as twice the absolute value of the difference between the transaction price and the quote midpoint. The third measure is the time-weighted average number of shares displayed in the book up to 10 cents from the best posted prices in a 10-minute interval. The short-term volatility measure is defined as the highest mid-quote in a 10-minute interval minus the lowest mid-quote in the same interval, divided by the midpoint between the high and the low.

We find that, in general, connected stocks have lower bid-ask and effective spreads, lower short-term volatility, and higher market depth than non-connected stocks. The market quality has improved after the connect program, except for short-term volatility. Compared to non-connected stocks with similar stock characteristics, the displayed market liquidity of connected stocks, as measured by bid-ask spreads and market depth, has improved after the connect program. However, the actual trading costs of investors, as measured by the effective spreads, have increased after the connect program. This implies that market liberalization

significantly increases trading costs for Chinese domestic investors. The short-term volatility of connected stocks has also increased following the introduction of SHHKConnect, which indicates an increase in cross-market arbitrage activities. The increased effective spreads show that the adverse selection channel dominates in China's stock market. As a result, Chinese domestic investors indeed pay higher trading costs after opening the market to foreign investors.

Our paper's main contribution is to empirically test the theoretical ambiguity and shed light on whether market liberalization enhances or impedes market quality. To the best of our knowledge, this paper is the first in the literature to provide evidence for both the competition and adverse selection channel and the net effect of the two in the context of China's stock markets. Our paper uncovers an adverse consequence of market openness—an increase in the trading costs for investors.

This paper contributes to a rich empirical body of literature that examines the effect of high frequency trading on market quality using intra-day order level data. Hasbrouck and Saar (2013) propose a new measure of high frequency trading and use this measure to study how low-latency activity affects market quality both during normal market conditions and during a period of declining prices and heightened economic uncertainty. They find that low-latency activity improves traditional market quality measures — decreasing spreads, increasing displayed depth in the limit order book, and lowering short-term volatility. Jørgensen et al. (2017), Malinova et al. (2016), and Friederich and Payne (2015) study the impact on market liquidity from the introduction of a penalty for high order-to-trade ratios in Norway, Canada, and Italy, respectively. Malinova et al. (2016) and Friederich and Payne (2015) find that the policy is associated with a drop in market liquidity. However, Jørgensen et al. (2017) find that market quality, measured by depth, spreads, and realized volatility, remains largely unaffected.

This paper is also related to a few papers that study the effect of SHHKConnect on China's stock market. Many of these studies focus on asset pricing (Hui and Chan (2018), Liu et al. (2016) and Burdekin and Siklos (2018)), risk sharing (Chan and Kwok (2017)), price discovery (Sohn and Jiang (2016)), asymmetric impacts on Shanghai and Hong Kong stock markets (Bai and Chow (2017)), and volatility spillover (Lin (2017), Zhang and Jaffry (2015)

and Huo and Ahmed (2017)). Our paper complements the above literature by studying the impact of the SHHKConnect on high frequency market quality and the trading costs of investors on the SSE. We aim to empirically test the theoretically ambiguous effect of market openness and disentangle which channel (competition or adverse selection) dominates in China's stock market. We are also the first paper to examine the effect of SHHKConnect by using the high frequency intra-day data set to show a finer picture of changes in market quality with high resolution.

The remainder of the paper is organized as follows: Section 2 provides an overview of institutional background, Section 3 develops the hypotheses, Section 4 describes the data and our empirical approach, Section 5 represents our estimation results, Section 6 conducts some further tests and Section 7 concludes. The appendix contains a list of key variables in the paper.

2 Institutional background

China's stock market had over 3400 firms listed and $8.5 trillion in market capitalization in October 2017, which represents over 10% of the global stock market. International investors have increasing interest in investing in China because China's stock market offers high average returns and low correlations with other equity markets (Carpenter et al. (2017)). The Chinese government has made a sequence of policies to open its capital market to foreign investors with the hope that they will bring mature investment strategies and business models to promote healthy competition and benefit the long-term development of China's capital market. The China Securities Regulatory Commission (CSRC) approved the Qualified Foreign Institutional Investors program in 2002, launched SHHKConnect in 2014 and the Shenzhen-Hong Kong Connect program in 2016, and initiated the Shanghai-London Stock Connect program in 2015, which is expected to take effect in 2019.

The SHHKConnect cross-boundary investment channel was launched and commenced operation on November 17, 2014.[2] SHHKConnect creates mutual stock market access to

[2]Before SHHKConnect, the opening of the local stock market to international investors was the launch of QFII (Qualified Foreign Institutional Investors scheme in 2002, allowing licensed foreign investors to participate in the Chinese stock market with an initial quota of USD20 billion. Because the SHHKConnect program has less restrictions on trading and is granted a larger initial quota (about USD80 billion), it is considered a further step in the liberalization of China's stock market after the QFII scheme.

trading designated stocks listed on either the SSE or the Hong Kong Stock Exchange (SEHK). This new investment channel will enable investors in Hong Kong and mainland China to trade a specified range of listed stocks in each other's market through their respective local securities companies, thereby helping to promote the openness of China's capital markets.

Among the 1,018 stocks that are listed on the SSE, investors in Hong Kong can invest in 564 of them; this is referred to as 'northbound trading'. This sample of firms represents approximately 90% of the total market capitalisation of the SSE. On the other hand, mainland Chinese investors can invest in 263 SEHK-listed stocks, of the possible 1,789 stocks that are listed on the SEHK. Otherwise known as 'southbound trading', this represents approximately 80% of the market capitalisation of SEHK.

In general, the 564 eligible SSE-listed stocks that can be traded under SHHKConnect include all the constituent stocks of the SSE 180 Index and the SSE 380 Index, as well as A-shares that have corresponding H-shares cross-listed on the SEHK (but not included in the indices mentioned). The 263 eligible SEHK-listed stocks to be traded under SHHKConnect include all the constituents of the Hang Seng Composite LargeCap Index and Hang Seng Composite MidCap Index, as well as all the H-shares.

Mainland Chinese investors, who have an aggregate amount of CNY500,000 (i.e. USD 80,514) or more in their security and cash accounts with brokers, are eligible to invest in the SEHK through SHHKConnect. SHHKConnect has provided mainland Chinese investors with greater and easier access to the Hong Kong stock market, whereas previously, mainland Chinese investors had only a limited ability to invest in the SEHK directly. They may have done so by opening a trading account with a Hong Kong-based broker; however, mainland investors are subject to various constraints regarding funds flow in and out of China.

Although overseas institutional investors were able to invest in the SSE by acquiring Qualified Foreign Institutional Investor (QFII) licenses and QFII quotas prior to SHHK-Connect, the program offers much greater freedom for international investors to invest in China. Moreover, SHHKConnect offers an unprecedented opportunity for international retail investors to access the historically closed Chinese capital market. Instead of purchasing ETF products that invest in Chinese securities, or investing in mutual funds via their brokers, retail foreign investors can directly select and hold stocks listed on the SSE.

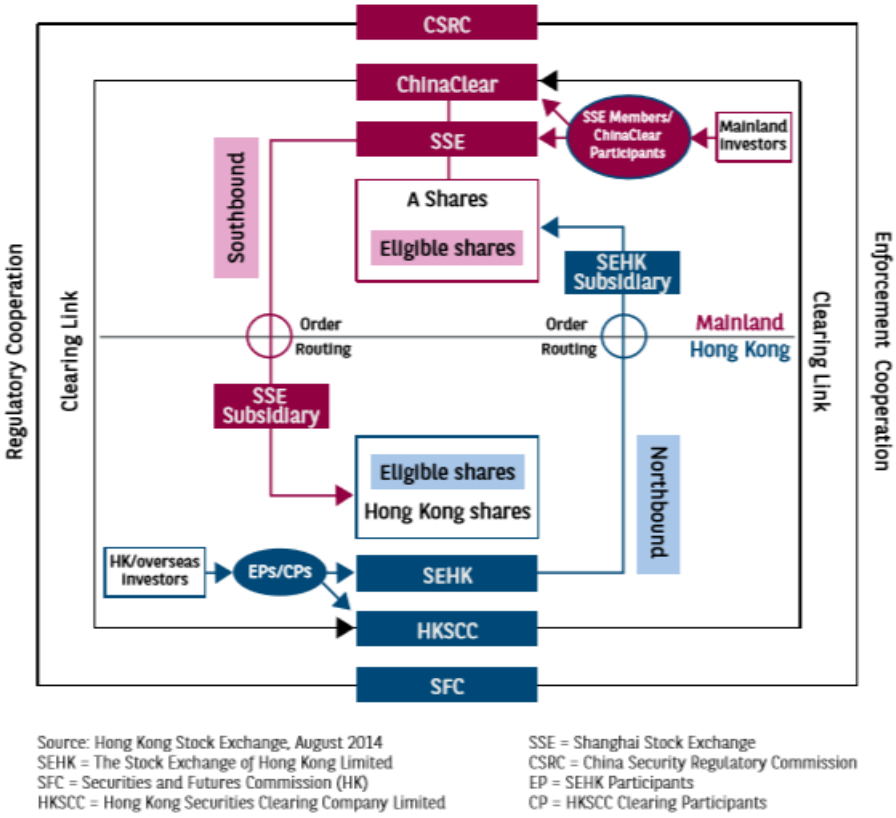


Figure 1: SHHKConnect Illustration

Figure 1 shows an overview of the regulatory cooperation for trading and clearing through SHHKConnect. Under SHHKConnect, the SSE and the SEHK established two subsidiaries, namely, SSE Subsidiary and SEHK Subsidiary, to act as non-member trading participants in the other market. The function of the subsidiaries is to facilitate cross-boundary order-routing for exchange participants in their home market. For example, the SEHK Subsidiary is established and located at the SSE as a local trading participant. The SEHK Subsidiary receives orders to trade stocks listed in China from exchange participants who are registered with the SEHK. It then routes the orders received in the trading system to the SSE for matching and execution. Similar arrangements are made by the SSE Subsidiary. Trading activities in both directions are limited to secondary market trading only; that is, investors cannot participate in initial public offerings (IPOs) across markets.

Clearing and settlement under SHHKConnect is conducted by the China Securities De-

pository and Clearing Corporation Limited (ChinaClear) and the Hong Kong Securities Clearing Company Limited (HKSCC). ChinaClear and HKSCC established a clearing link whereby the two clearing houses act as participants in each other. Under SHHKConnect, in either direction, securities are traded in local currency but settled in CNY. For instance, for southbound trades, Chinese investors will trade SEHK listed stocks in Hong Kong dollars. These trades will be settled with ChinaClear or its clearing participants in CNY. For the northbound trades, HKSCC will settle such trades with its clearing participants and ChinaClear in CNY. This implies that all currency conversions are effected outside China, a process strategically supporting the Chinese government in internationalizing the Chinese CNY. The stock and money settlements in each direction follow the clearing and settlement cycles in the other market. That is, the northbound trades are settled following settlement rules in the SSE, which is T day for stock settlement and T+1 for money settlement, and vice versa for the southbound trades.

During the sample period from October to December 2014, quotas were imposed for each trading direction (i.e. north- and south-bound). The trades are subject to a maximum cross-boundary investment quota, namely, Aggregate Quota, as well as the Daily Quota. The quotas aim to cap the amount of funds flowing in and out of mainland China. The China Securities Regulatory Commission increased the daily southbound and northbound quotas for SHHKConnect on May 1, 2018. The southbound quota has risen to 42 billion CNY from 10.5 billion, and the northbound quota has risen to 52 billion yuan from 13 billion since then. Purchasing activities through SHHKConnect will be suspended when either quota is reached. Sell orders are always allowed regardless of quota level. The two exchanges distribute market data regarding respective trading quotas free of charge. The SSE updates the daily quota balance for southbound trading every sixty seconds and SEHK updates the real-time daily quota balance for northbound trading every five seconds.

3 Hypotheses development

The SHHKConnect program allows foreign investors to trade and compete directly with Chinese domestic investors on the SSE. Opening the market to foreign investors increases competition for liquidity provision. Theoretical models of liquidity provision argue that bid-

ask spreads decrease and market depth increases as the level of competition increases (Ho and Stoll (1983), Dutta and Madhavan (1997) and Brogaard and Garriott (2014)). We thus lay out our first hypothesis. SSE is a limit order market that keeps record of an order book containing the list of interested buyers and the list of interested sellers. In limit order books, traders either supply liquidity by posting non-marketable limit orders that specify a price and a total order size, or demand liquidity by submitting market orders that yield immediate execution. The limit orders are matched with market orders following a price/time priority rule—orders are first ranked according to their prices; orders of the same price are then ranked depending on when they were entered. Limit orders are potentially executed at better prices than market orders, but they run the risk of non-execution and are exposed to being adversely selected if the security's value moves past the limit price before the limit order can be cancelled.

Investors who place limit orders compete against each other by quoting better prices and undercutting existing limit orders to gain price priority for execution. When the limit order book is open to foreign investors, liquidity suppliers compete more fiercely and, ceteris paribus, submit more aggressive orders. As competition intensifies, we expect to observe narrower bid-ask spreads and more depth in the limit order book. We thus lay out our first hypothesis.

Hypothesis 1: After the implementation of the SHHKConnect program, connected stocks experience significant lower bid-ask spreads and higher market depth than non-connected stocks with similar characteristics because of competition from foreign investors.

The second hypothesis is regarding the adverse selection channel. When the market is open to more sophisticated foreign investors, we expect to see more cross market arbitrage activities that increase adverse selection and speed arbitrage, therefore putting upward pressure on bid-ask spreads (Glosten and Milgrom (1985), Foucault (1999), Foucault et al. (2017) and Biais et al. (2015)). Cross market arbitrage is the practice of exploiting disparities in the price at which equivalent goods can be traded in different markets. Such disparities can arise in financial markets because of market fragmentation, latency variations, and information asymmetry. Because cross market arbitrage opportunities disappear quickly, investors using such strategies usually submit market orders instead of limit orders to get immediate

execution. Limit order traders face an adverse selection risk, as arbitragers are privately better informed about the price movement of the security in the next instant. As a result, limit order traders quote less aggressively to avoid the risk of being picked off by the better informed arbitragers. Hence, bid-ask spreads widen as cross market arbitrage becomes more active.

High frequency arbitrage activities are also associated with intensive order submissions, updates and cancellations, that increase short-term volatility (Hasbrouck (2018) and Zhang (2010)). The empirical literature in high frequency trading has documented an increase in the number of fleeting limit orders, which are placed and then immediately cancelled without execution. We also observe a shorter resting time before a limit order is cancelled. More mini-crashes in individual securities are observed (Menkveld (2018)), in which there are substantial increases or decreases in prices as liquidity disappears and market orders result in dramatic price changes within milliseconds. Therefore, we expect to see higher short-term volatility. This leads to our second hypothesis.

Hypothesis 2: Connected stocks will have higher bid-ask spreads and higher short-term volatility after the connect program than non-connected stocks with similar characteristics due to the increase in high frequency cross-market arbitrage activities.

Hence, opening the stock market to more sophisticated foreign investors has ambiguous effects on bid-ask spreads, depending on which mechanism dominates. Our paper aims to empirically investigate this theoretical ambiguity and disentangle which channel dominates in China's stock market. We also show the net effect of the competition and adverse selection channels by examining the influence of market openness on the trading costs of investors on SSE.

4 Data and sample

4.1 Data

Our primary data source is the SSE high frequency order-level data with millisecond timestamps in the CSMAR database provided to us by GTA Information Technology. The SSE operates an electronic limit order book with price and time execution priority. The CSMAR database provides real-time information about quotes with prices and quantities in the first

ten levels of the LOB's ask- and bid-sides, order queues that show the time priority of orders, and transaction data with transaction price and quantity on the SSE. The data are comprised of time-sequenced snapshots that describe the history of trade and limit order book activity. As soon as there is a change in price, quantity, or order queue at any level of the book due to a newly placed, cancelled (or partially cancelled), or executed (or partially executed) order, a new snapshot (identified by a unique message ID) of the entire book is created. These data provide a detailed picture of the trading process and the state of the SSE limit order book with millisecond timestamps.

In this paper, we only focus on the effect of northbound trading on the SSE for three reasons. First, as the largest developing stock market in emerging economies, the SSE is not as mature a market as the SEHK. Hence, we expect to see a stronger effect of opening the market to foreign investors on the market quality of the SSE. The SEHK is already an open and mature market, so we expect to see little effect of the SHHKConnect program on the trading environment at the SEHK. Second, northbound trading is more active than southbound trading. Based on the historical data reported on SEHK, in September 2018, the average daily turnover for northbound and southbound trading is 10.45 billion RMB and 5.79 billion HKD, respectively. Northbound trading also has a higher daily quota than southbound trading. Third, southbound trading requires the investor to have a minimum of 0.5 million RMB to participate in SHHKConnect. However, there are no such restrictions for northbound trading.

4.2 Sample and summary statistics

Our sample is constructed to capture the variations in market quality for the connected stocks around SHHKConnect, taking the non-connected stocks as control stocks. We identify all the connected and non-connected stocks that are SSE listed from October to December 2014. Since SHHKConnect was launched on November 17, 2014, we divided the sample into two sub-samples, namely a pre-connect sub-sample from Oct 1 to Nov 16, 2014, and an after-connect sub-sample from November 17 to Dec 31, 2014. The SSE composite index experienced rapid growth during that time, with the SSE composite index starting the period at 2,363 and ending it at 3,234. Most of the growth occurred after the SHHKConnect program

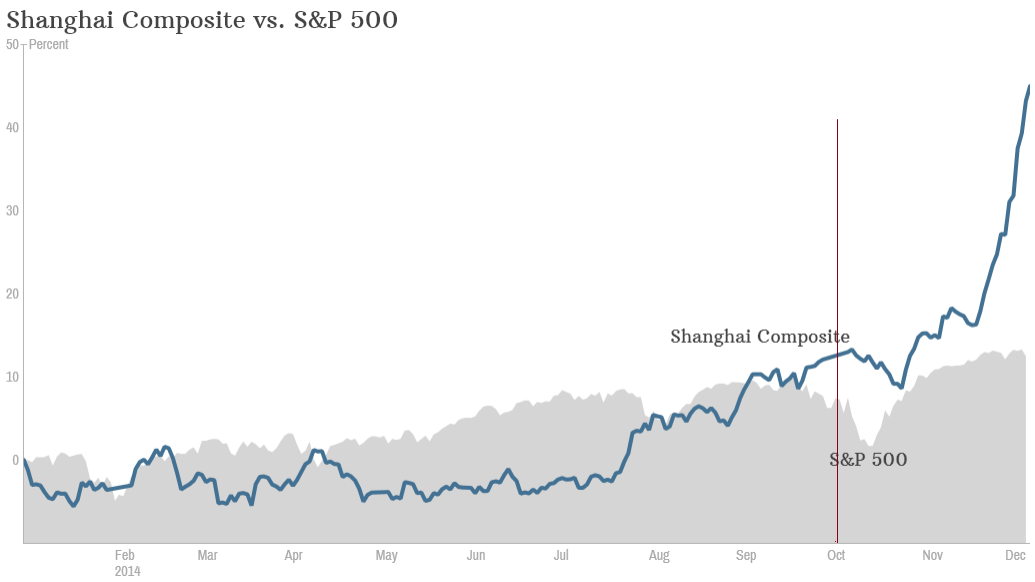


Figure 2: Shanghai Composite vs. S&P 500

with the index being at 2,475 on Nov 17, 2014. Figure 2 shows that the SSE composite index return is roughly three times the gain made by the S&P 500 in 2014.

We construct summary statistics over 10-minute intervals. There are 61 trading days during our sample period, and each normal trading day has 24 intervals. We start with 564 connected stocks and 454 non-connected stocks listed on the SSE. We eliminate all firm-interval pairs with trading suspensions on either the buy or sell side of that stock during the interval. We only consider the trading hours with continuous auctions (9:30 to 11:30 and 13:00 to 15:00) by excluding opening and closing auctions. Net of these exclusions, the sample contains 765,254 connected stock-interval pairs and 587,957 non-connected stock-interval pairs. To minimize the effect of outliers, we winsorize all variables at the top and bottom 1% of each variable's distribution.

Table 1 provides summary statistics for the connected and non-connected stocks during the sample period. Panel A summarizes market capitalization, turnover, volume, and realised volatility. For the connected stocks, market capitalization ranges from \$2.766 million to \$115 million, with a median of slightly over \$13 million. The sample also spans a range of trading activity and volatility levels. The most active stock exhibits an average volume of 26.25 million shares; the median is about 9.11 million shares. Realised volatility within each interval ranges from 0 to 5.528%, with a median of 0.215%. Panel B summarizes interval

average quoted spread, effective spread, short term volatility, market depth, and average trade size for both connected and non-connected stocks during the sample period. For the connected stocks, the average bid-ask spread is 13.32 bps, and the average effective spread is 47.14 bps.

4.3 Empirical specification

By means of a difference-in-differences analysis, we study the effect of market openness on the four market quality measures defined in Section 4. This methodology compares the market quality of a sample of treatment stocks that are eligible for SHHKConnect to the market quality of a matched sample of control stocks that are not eligible, but are otherwise comparable, before and after policy changes that cause an exogenous shock to market openness. Difference-in-differences estimation combines a control group with the treated sample to "difference out" confounding factors and isolates the effect of an event. We use the difference-in-differences approach to determine the effect of a change in market openness on market quality.

We start by identifying a large exogenous shock to stock market openness—SHHKConnect. Prior to SHHKConnect, there are various restrictions for foreign investors to trade directly in China's stock market. The SHHKConnect program allows foreign investors who have an account in Hong Kong to trade a specified group of eligible stocks directly on the SSE. Examination of the change in market quality following the change in market openness due to SHHKConnect provides a natural experiment for our analysis. Among the 1,018 stocks that are listed on the SSE, 564 of them are eligible to trade through northbound trading and the rest are not eligible. We make use of this cross-section variation in market openness between eligible stocks and non-eligible stocks to study the influence of the policy change. We use the 564 connected stocks that are directly affected by the SHHKConnect program as the treated sample, and the rest of the stocks that are not directly affected by the connection as the controlled sample.

Connected stocks in general have higher market capitalization, higher trading volume and better liquidity (see Table 1). To address the selection bias, we construct a treatment group and a control group of stocks using the propensity score matching method. We match the 564

Table 1: Summary Statistics

Panel A Stock characteristics

	Connected				Non-Connected			
	MarketCap (¥million)	Turnover (¥million)	Volume (million)	Realised Volatility	MarketCap (¥million)	Turnover (¥million)	Volume (million)	Realised Volatility
Mean	21.39	15.4	1.637	0.44	5.42	4.745	0.5265	0.422
Median	13.4	7.095	0.6393	0.226	3.712	2.177	0.22	0.204
Std Dev	21	24.5	2.973	0.672	8.953	10.2	1.163	0.696
Max	115	164	20.3	5.528	115	164	20.3	5.528
Min	2.766	0.0027	0.0032	0	0.0828	0.0027	0.0032	0

Panel B Market liquidity and activity measures

	Connected					Non-Connected				
	BidAskSpd (bps)	EffSpd (bps)	HighLow (bps)	Depth (million)	Trade size (shares)	BidAskSpd (bps)	EffSpd (bps)	HighLow (bps)	Depth (million)	TradeSize (shares)
Mean	13.32	47.14	65.31	10.4	2265	16.49	117.18	60.82	3.593	2140
Median	11.75	17.11	48.31	5.091	1967	14.64	42.24	45.42	1.999	1924
Std Dev	6.964	88.62	57.14	16.2	1303	8.675	194.58	54.15	5.124	1195
Max	55.99	1193.4	369.4	106	7760	55.99	1193.4	369.4	81.9	7760
Min	3.565	0.1208	0	0.2165	336.2	3.565	0.1208	0	0.2165	336.2

Notes: The sample consists of 450 connected stocks and 450 matched non-connected stocks over the period from Oct 2014 to December 2014 (61 trading days). A firm-interval pair is dropped from the sample if there are trading suspensions on either the buy or sell side of that stock during the interval. The opening and closing auctions are excluded from the sample. All variables are winsorized at the top and bottom 1% of each variable's distribution. Panel A reports interval average statistics for turnover, trading volume, and realised volatility. Market capitalization is as of the end of December 2014. Panel B reports interval average statistics on market quality and activity measures. Market depth and bid-ask spread are time-weighted averages for each firm during each interval. The effective spread is defined as twice the absolute value of the difference between the transaction price and the quote midpoint, and the average is value-weighted. The short-term volatility (HighLow) is defined as the highest midquote in an interval minus the lowest midquote in the same interval, divided by the midpoint between the high and the low, expressed in basis points. Trade size is the equally weighted average shares per transaction in each interval.

connected stocks with the 454 non-connected stocks according to four firm characteristics: market capitalization, book-to-market ratio, return-on-assets, and total volatility at the end of December 2014. We then find each connected stock a matched non-connected control stock using the nearest neighbour matching technique. The matching is conducted using R package "MatchIt" (Ho et al. (2007)). We match each connected stock with one non-connected stock using a propensity score, except that 4 non-connected stocks are discarded as their propensity scores fall outside the support of the distance measure (i.e. removed as extreme values). This procedure results in a final sample of 450 connected stocks with valid non-connected control stocks.

Table 2 summarizes the matching results for the treatment group and the control group before and after the match. We report Wilcoxon test statistics instead of the t-statistics because the distribution is not normal. When data does not follow a normal distribution, the results of t-statistics can be misleading, and the nonparametric Wilcoxon test is a more appropriate choice, since it does not have an assumption about the underlying distribution. We find the difference between the mean of non-connected and connected stocks has decreased for all variables except volatility. The data is heavily skewed, hence no improvement in the difference of the mean in one variable does not mean it is bad. The Wilcoxon statistics of all variables are actually decreased after matching. By comparing the summary statistics of propensity scores before and after matching, we can see that all of them improved after matching (i.e. propensity scores of connected and non-connected stocks got closer).

We then construct a set of market quality and activity metrics to be used as dependent variables in the difference-in-differences estimation. For each stock, we use high frequency intra-day data with millisecond timestamps to construct key variables over 10-minute intervals. We build four market quality measures and three market activity measures as dependent variables in the difference-in-differences regressions.

Following Hasbrouck and Saar (2013), we use four measures of market quality, namely bid-ask spread, effective spread, market depth, and short-term volatility. The first measure is the time-weighted average quoted spread (best ask price minus best bid price) on the SSE in an interval. The second measure is the value-weighted average effective spread (or total price impact) of all trades on the SSE during the 10-minute interval, where the effective

Table 2: Propensity Score Matching Results

Panel A	Summary of Matching Variables Before and After Match							
	Before Match				After Match			
	Non-connected	Connected	Difference	Wilcox	Non-connected	Connected	Difference	Wilcox
MktCap (million)	5.41	38.98	-33.58	34660	5.42	21.39	-15.97	20834
Volatility	0.03	0.03	0.00	153142	0.03	0.03	0.00	126211
ROA	-0.11	0.04	-0.15	82473	0.01	0.05	-0.04	58200
BM	0.98	1.48	-0.50	95649	0.99	1.28	-0.29	77333

Panel B	Propensity Score Matching Results							
	No. of Obs.	Min	P5	P50	Mean	SD	P95	Max
Non-connected.before	454	-484.30	-2.21	-0.71	-1.74	22.83	1.07	18.37
Connected.before	564	-6.05	-0.74	0.96	5.03	18.67	21.19	263.02
Non-connected.after	450	-5.15	-2.07	-0.70	-0.56	1.57	1.09	18.37
Connect.after	450	-0.31	-0.17	1.15	2.45	3.26	9.37	18.01

Notes: Panel A summarizes the univariate comparisons between the treatment and control stocks' characteristics before and after the match and their corresponding Wilcoxon statistics. The first two columns are the means of non-connected and connected stocks for the four matching variables. The third column shows the difference. The 4th columns is the Wilcoxon test statistics. Larger values indicate bigger differences between the treatment group and the control group. The four matching variables are market capitalization (in millions), book-to-market ratio, return-on-assets, and total volatility at the end of December 2014. Panel B reports the distribution of estimated propensity scores for the treatment stocks and control stocks before and after the match.

spread is defined as twice the absolute value of the difference between the transaction price and the quote midpoint (the quote midpoint is calculated from the best bid and ask prices observed immediately prior to the trade). Both the bid-ask spread and the effective spread are expressed in basis points of the mid-quote. The third measure is the time-weighted average number of shares in the book up to 10 cents from the best posted prices. The short-term volatility measure is defined as the highest mid-quote in an interval minus the lowest mid-quote in the same interval, divided by the midpoint between the high and the low (and multiplied by 10,000 to express it in basis points). We next explore some market activity measures to understand possible mechanisms for the changes in market quality. Following Friederich and Payne (2015), the market activity measures are the number of trades, the average trade size, and the average turnover per trade in each interval.

We regress these dependent variables on a set of treatment indicators that includes a dummy variable picking out the connected stocks on the SSE (Connected), a time dummy picking out the period after the SHHKConnect implementation (Policy), and the interaction of those two dummies. If there is any difference in the behaviour of the variable for the connected and control sample stocks after the SHHKConnect introduction, it will appear as a significant coefficient on the interaction variable (Connected $\times$ Policy). The coefficient β_3 in the following equation shows the effect of the policy change. Denoting the dependent variables of interest with $y_{i,t}$, we estimate the following regression equation with the matched sample

$$y_{i,t} = \beta_0 + \beta_1 \times Connected_i + \beta_2 \times Policy_t + \beta_3 \times Connected_i \times Policy_t + e_{it} \qquad (1)$$

where the dependent variable y represents one of the market quality and activity measures. *Connected* is a dummy variable that equals one for connected stocks and zero for non-connected stocks. *Policy* is a time dummy that equals one for after the SHHKConnect program and zero for before the program.

We examine the sensitivity of the analysis to the presence of control variables to allow for the possibility that these control variables might affect the change in market quality. We considered three control variables: realised volatility, turnover, and market capitalization.

In particular, we estimated the following regression equation

$$y_{i,t} = \beta_0 + \beta_1 \times Connected_i + \beta_2 \times Policy_t + \beta_3 \times Connected_i \times Policy_t + \alpha \times X_{it} + e_{it} \quad (2)$$

where y, *Connected*, and *Policy* are defined the same as in Equation 1. X is a vector of control variables, including realised volatility, turnover and market capitalization. Realised volatility and turnover are calculated for each interval and stock; thus, they vary across both stocks and time. Market capitalization is the market value of all stocks at the end of 2014. When we use an activity measure as a dependent variable in this regression, the control variables are realised volatility and market capitalization only.

As a robustness test, we also estimate the difference-in-differences regression with stock fixed effects and interval time fixed effects to account for any unobserved time-invariant characteristics of individual stocks, and at the same time allowing for stock-invariant time fixed effects. Specifically, we estimate the following regression equation

$$y_{i,t} = \alpha_i + \lambda_t + \beta_1 \times Connected_i + \beta_2 \times Policy_t + \beta_3 \times Connected_i \times Policy_t + e_{it} \quad (3)$$

where y, *Connected*, and *Policy* are defined the same as in Equation 1. α_i captures stock fixed effects and λ_t allows for time fixed effects.

5 Empirical results

5.1 Market quality

Estimation results for the four dependent variables measuring market quality appear in Table 3 and 4. First, we note that the SHHKConnect introduction has a significant effect on all four market quality measures. In general, we find that after the SHHKConnect event, the bid-ask spreads and effective spreads decrease for both connected and non-connected stocks — the estimated coefficient on the time dummy β_2 is significantly negative for bid-ask spreads and effective spreads. Furthermore, the market depth and short-term volatility both increase after the SHHKConnect event— the estimated coefficient on the time dummy β_2 is significantly positive for market depth and short-term volatility. The significant negative coefficients β_1 for bid-ask spreads, effective spreads, and short-term volatility, and the sig-

Table 3: Difference-in-differences regression results for market quality measures

	Spread	EffSprd	HighLow	Depth	Spread	EffSprd	HighLow	Depth
Connected	-16.20***	-138.0***	17.05***	1.764***	-2.882***	-36.10***	-3.081***	0.868***
	(-46.94)	(-30.41)	(10.63)	(150.84)	(-130.41)	(-170.00)	(-24.08)	(301.62)
Policy	1.988***	9.673***	11.81***	-0.124***	-0.274***	-5.857***	7.782***	0.0792***
	(33.76)	(13.07)	(24.43)	(-30.44)	(-11.97)	(-22.79)	(55.54)	(29.78)
Connected_policy	-0.717***	0.642**	15.44***	0.110***	-0.551***	1.976***	14.15***	0.105***
	(-37.84)	(2.51)	(83.01)	(80.7)	(-18.93)	(7.03)	(71.27)	(26.56)
Constant	27.62***	158.9***	23.84***	13.32***	16.64***	61.76***	56.60***	14.51***
	(80.73)	(36.04)	(24.91)	(1534.43)	(978.8)	(317.27)	(595.14)	(7499.51)
Firm fixed effects	yes	yes	yes	yes	no	no	no	no
Time fixed effects	yes	yes	yes	yes	no	no	no	no
N	1191607	1191607	1191607	1191607	1191607	1191607	1191607	1191607
R^2	0.605	0.24	0.182	0.906	0.0408	0.0526	0.0237	0.157

Notes: This table reports the results of 10-minute interval panel difference-in-differences estimation of variables measuring bid-ask spreads (Spread), effective spreads (EffSprd), short-term volatility (HighLow) and market depth (Depth) for the connected and non-connected sample. Two indicator variables pick out connected sample stocks and the period after SHHKConnect policy, respectively, and a further indicator variable interacts with the previous two. All variables, stock fixed effects and time fixed effects are defined in Section 4.3. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

Table 4: Difference-in-differences regression results for market quality measures with control variables

	Spread	EffSprd	HighLow	Depth	Spread	EffSprd	HighLow	Depth
Connected	-0.391***	-2.177***	-2.435***	0.0946***	0.474	3.386	7.014	-0.0549
	(-18.93)	(-10.49)	(-26.02)	(33.6)	(.)	(.)	(.)	(.)
Policy	-0.191***	-2.274***	0.398***	0.0254***	0.981***	6.098***	12.35***	-0.0698***
	(-10.19)	(-9.89)	(4.68)	(11.1)	(18.87)	(8.31)	(41.52)	(-17.92)
Connected_policy	-0.508***	5.937***	4.024***	0.0532***	-0.709***	3.802***	2.606***	0.0849***
	(-20.25)	(22.99)	(32.88)	(15.97)	(-41.37)	(14.75)	(22.51)	(63.94)
RV	5.596***	7.214***	51.34***	-0.0817***	3.995***	2.831***	44.40***	-0.128***
	(352.75)	(70.13)	(460.72)	(-54.70)	(3.995)***	(2.831)***	(44.40)***	(-0.128)***
lnTurnOver	-2.424***	-17.19***	13.55***	0.232***	-1.523***	-10.35***	21.06***	0.122***
	(-318.12)	(-189.70)	(366.41)	(270.93)	(-186.41)	(-78.00)	(380.09)	(199.01)
lnMarketCap	0.743***	-6.016***	-11.04***	0.282***	-2.016	-20.21	-27.93	0.278
	(70.23)	(-55.30)	(-228.63)	(207.9)	(.)	(.)	(.)	(.)
Constant	34.24***	419.3***	61.63***	5.747***	75.64	598.3	289.2	7.488
	(195.92)	(205.6)	(97.12)	(286.98)	(.)	(.)	(.)	(.)
Firm fixed effects	no	no	no	no	yes	yes	yes	yes
Time fixed effects	no	no	no	no	yes	yes	yes	yes
N	1191607	1191607	1191607	1191607	1191607	1191607	1191607	1191607
R^2	0.295	0.21	0.63	0.409	0.684	0.25	0.682	0.912

Notes: This table reports the results of 10-minute interval panel difference-in-differences estimation with control variables. The dependent variables are bid-ask spreads (Spread), effective spreads (EffSprd), short-term volatility (HighLow), and market depth (Depth) for the connected and non-connected sample. Two indicator variables pick out connected sample stocks and the period after SHHKConnect policy, respectively, and a further indicator variable interacts with the previous two. The three control variables, realised volatility (RV), turnover (lnTurnOver), and market cap (lnMarketCap) are defined in Section 4.3. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

nificant positive coefficient for market depth, indicate that connected stocks in general have narrower bid-ask and effective spreads, lower short-term volatility, and higher market depth than non-connected stocks. To sum up, the analysis shows that the market quality is better for the connected stocks than for the non-connected stocks. Three out of four market quality measures have improved after the SHHKConnect implementation. The estimated coefficient on the dummy variable identifying connected stocks became insignificant after adding the control variables and fixed effects. The sign of the coefficient on the policy time dummy changes direction for bid-ask spread, effective spread and market depth after controlling for realized volatility, turnover, market cap and fixed effects.

Second, the estimated coefficients on the key interaction variable β_3 are all highly significant. Starting with the bid-ask spreads and market depth, the dummy coefficient β_3 shows that the bid-ask spreads decrease significantly and the market depth increases significantly for the connected stocks relative to non-connected stocks after the SHHKConnect. The improved displayed liquidity is consistent with increased competition between the two connected stock markets. The sign of the coefficient on the interaction term remains the same even after controlling for realized volatility, turnover, market cap and fixed effects.

In contrast to the increase in displayed liquidity, the analysis of effective spreads and short-term volatility shows statistically significant positive effects on the connected stocks after the SHHKConnect: the coefficient β_3 on the interaction dummies for effective spreads is positive and significant. The sign of the coefficient does not change after controlling for realized volatility, turnover, market cap and fixed effects. Unlike the bid-ask spread, which measures the cost of a small round-trip transaction, the effective spread reflects the true trading costs obtained by investors.[3] The effective half spread is defined as the difference between the price at which a market order executes and the mid-quote observed in the limit order book the instant before. The magnitude of the difference-in-differences estimator on effective spreads suggests that, on average, opening the market results in an increase of approximately 4 bps in the effective spread for connected stock than for the non-connected stocks. Although the change in magnitude seems to be small, if we multiply the 4 basis points increase by the average turnover per day for all connected stocks in our sample, the

[3]The transaction has to be small enough that it can be filled at the best bid and ask prices

increase in trading costs amount to 0.9 million CNY per day.

Two findings emerge from our results. First, market openness improves displayed liquidity, which is consistent with our competition hypothesis. Second, and more importantly, market liberalization is associated with increased trading costs for investors despite the enhanced displayed liquidity in the limit order book. The increase in effective spreads can be explained by an increase in trade size and higher order cancellation rate. We examine how trade size changes after the connect program in Subsection 5.2.

Limit order cancellation is well studied in the high frequency trading literature. Short-lived limit orders that are only on the limit order book for a few seconds before being cancelled are called fleeting orders. Hasbrouck and Saar (2009) document a tremendous large portion and fast speed of limit order cancellation associated with high frequency trading. They show that 90% of orders on INET are limit orders. Among all the limit orders, 93% of them are cancelled or revised. For the cancelled orders, more than one third of them are cancelled within 2 seconds after submission. Hence, displayed liquidity is not a good estimation for liquidity available to investors. Fleeting orders often arise from increasing high frequency trading activities. This is consistent with our second hypothesis that opening the market will attract more foreign quantitative traders, whose trading strategies require the submission of a lot of fleeting orders. The increase in trading costs can also be explained by the increase in cross-market arbitrage activities that intensify adverse selection. The observed higher short-term volatility is also expected and aligned with our second hypothesis that there are more cross-market arbitrage activities after the stock markets are connected.

We next examine the sensitivity of this association to the presence of control variables. To allow for the possibility that volatility factors, turnover, and market capitalization might drive market quality measures, we include these control variables in the regression. The results with control variables are reported in Table 4. The sign and the significance of the coefficients β_3 on all four market quality measures remain the same. The magnitude of these coefficients becomes smaller because the control variables explain some of the changes affected by the policy. The coefficients on control variables are all significant and the signs are as expected. Greater turnover is associated with narrower bid-ask spreads and higher market depth. Stocks with larger market caps are more liquid, and they have smaller effective

spreads, lower short-term volatility and more depth. Higher realised volatility increases bid-ask and effective spreads, but decreases market depth. We also consider the sensitivity of the analysis to the inclusion of stock fixed effects and time fixed effects. We report the difference-in-differences regression results with fixed effects in Table 3 and 4. The sign and significance of the coefficients on the cross term remain the same for all four market quality measures.

5.2 Market activity

Estimation results for the three dependent variables measuring market activities appear in Table 5 and 6. The estimates from the regressions that use measures of market activity as dependent variables also tend to be consistent with our hypothesis that opening the market to more sophisticated foreign investors increases trading activities in the market. In general, connected stocks are more actively traded than non-connected stocks. We observe more transactions per interval, larger trade size, and higher turnover per transaction for connected stocks than for non-connected stocks. On average, trading becomes more active for all stocks after the introduction of the SHHKConnect.

The significant coefficients on the cross term show that connected stocks have more trades per interval, slightly larger trade size, and, thus, slightly higher turnover per trade. The increase in transactions and turnover are both statistically and economically significant, at around 23% and 0.06%, respectively. The increase in trade size is smaller in magnitude, at around 0.5%, and the sign of the coefficient changes after controlling for fixed effects and other control variables.

One unique feature of the Chinese stock market is that the average trade size is much larger than in the US equity markets. The average trade size on the New York Stock Exchange (NYSE) is about 200 shares per trade (Angel et al. (2011)), whereas the average trade size on the SSE is around 2200 shares per trade. As a result of the small trade size in the US market, the effective spread is equal to or less than the quoted spread.[4] However, the trade size in the Chinese market is much larger; hence, the effective spreads that investors pay are also much larger than the bid-ask spreads. In the Chinese market, it is quite possible that

[4]The effective spreads are less than the bid-ask spreads when there are price improvements or rebates.

Table 5: Difference-in-differences regression results for market activity

	Transactions	TradeSize	Turnover	Transactions	TradeSize	Turnover
Connected	1.007***	0.0548***	0.396***	3.146***	-0.0667***	2.591***
	(332.02)	(34.93)	(201.43)	(112.99)	(-3.57)	(156.43)
Policy	0.210***	-0.0382***	0.0263***	-0.230***	-0.0968***	-0.0797***
	(63.88)	(-25.25)	(11.02)	(-30.91)	(-27.07)	(-22.13)
Connected_policy	0.231***	0.00500**	0.0559***	0.272***	-0.0143***	0.0687***
	(54.64)	(2.38)	(21.43)	(105.74)	(-11.57)	(55.62)
Constant	4.592***	7.535***	9.527***	2.356***	7.589***	7.158***
	(1900.1)	(6757.91)	(5284.47)	(107.33)	(474.86)	(540.84)
Firm fixed effects	no	no	no	yes	yes	yes
Time fixed effects	no	no	no	yes	yes	yes
N	1191607	1191607	1191607	1191607	1191607	1191607
R^2	0.209	0.00353	0.0872	0.721	0.669	0.798

Notes: This table reports the results of 10-minute interval panel difference-in-differences estimation of variables measuring the number of transactions in each interval (Transactions), average trade size (TradeSize), and average turnover per transaction (Turnover) for the connected and non-connected sample. Two indicator variables pick out connected sample stocks and the period after SHHKConnect policy, respectively, and a further indicator variable interacts with the previous two. All variables are defined in Section 4.3. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

Table 6: Difference-in-differences regression results for market activity with control variables

	Transactions	TradeSize	Turnover	Transactions	TradeSize	Turnover
connected	0.151***	0.0442***	-0.236***	0.101	0.0921	-0.111
	(112.03)	(25.62)	(-140.64)	(.)	(.)	(.)
policy	0.0402***	-0.0686***	0.00539***	-0.00455	-0.0232***	0.00869***
	(30.92)	(-46.98)	(3.32)	(-1.51)	(-7.15)	(2.88)
connected_policy	0.0226***	-0.0364***	0.0423***	0.00879***	-0.0781***	0.0122***
	(13.53)	(-17.91)	(20.52)	(8.34)	(-70.86)	(11.69)
RV	0.134***	0.0567***	0.0135***	0.0567***	-0.0632***	-0.0572***
	(166)	(65.26)	(16.62)	(114.86)	(-122.39)	(-118.20)
lnturnOver	0.671***	0.116***		0.753***	0.211***	
	(1401.35)	(209.4)		(1681.57)	(466.00)	
lnmarketCap	-0.0513***	-0.0951***	0.352***	-0.314	-0.335	0.480
	(-78.11)	(-113.06)	(419.69)	(.)	(-0.02)	(.)
Constant	-3.931***	7.737***	1.736***	0.483	11.25	-3.396
	(-358.68)	(641.43)	(108.54)	(.)	(.)	(-0.01)
Firm fixed effects	no	no	no	yes	yes	yes
Time fixed effects	no	no	no	yes	yes	yes
N	1191607	1191607	1191607	1191607	1191607	1191607
R^2	0.878	0.0744	0.424	0.954	0.742	0.858

Notes: This table reports the results of 10-minute interval panel difference-in-difference estimation with control variables. The dependent variables are the number of transactions in each interval (Transactions), average trade size (TradeSize), and average turnover per transaction (Turnover) for the connected and non-connected sample. Two indicator variables pick out connected sample stocks and the period after SHHKConnect policy respectively, and a further indicator variable interacts with the previous two. The three control variables, realised volatility, turnover and market cap are defined in Section 4.3. The t-statistics are reported in parenthesis. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

when we observe narrower bid-ask spreads, the effective spreads do not necessarily decrease due to the large trade size per transaction.

5.3 Cross-listed stocks

In this section, we study the effect of SHHKConnect on the market quality of cross-listed stocks on both the SSE and the SEHK. Theoretical models of market fragmentation suggest that the effect is ambiguous (Baldauf and Mollner (2017)). Before SHHKConnect, investors in Shanghai could only trade the cross-listed stocks on the SSE. After the connect program, investors in Shanghai can now trade the cross-listed stocks on both the SSE and the SEHK. The competition for investors between the SSE and the SEHK places downward pressure on bid-ask spreads and trading costs of investors (Pagnotta and Philippon (2018) and Colliard and Foucault (2012)).

On the other hand, for cross-listed stocks, it is easier to conduct cross-market arbitrage trading after the connect program. Sohn and Jiang (2016) find that the SEHK contributes more to price discovery than the SSE for the cross-listed stocks. Mainland investors are concerned that when prices on the SEHK change before prices adjust on the SSE, informed foreign investors may race to the SSE to do cross-market arbitrage through northbound trading. Although the cross-market arbitrage keeps the prices in different markets from diverging without bound (Hasbrouck (1995)), increased cross-market arbitrage activities dampen market liquidity by increasing bid-ask spreads and trading costs of investors (Glosten and Milgrom (1985), Foucault (1999), Foucault et al. (2017), Biais et al. (2015), Hasbrouck (2018) and Zhang (2010)). We investigate this ambiguity empirically to see which mechanism dominates China's stock market after the connect program.

There are 71 stocks in our sample that are cross-listed on both the SSE and SEHK. The summary statistics of the cross-listed stocks are reported in Table 7. Out of the 71 cross-listed stocks, 70 stocks are included in SHHKConnect and only one cross-listed stock is not included. For the cross-listed subsample, we dropped the dummy variable indicating the difference between connected and non-connected stocks, because there is only one stock in the cross-listed subsample that is non-connected, which does not allow much variation between connected and non-connected stocks. The regression results for the cross-listed

Table 7: Summary statistics for cross-listed stocks

	Mean	Median	Std Dev	Max	Min
Market Cap (¥million)	42	36.6	25.1	103	7.203
TurnOver (¥million)	20.1	10.3	27.9	164	0.0027
Volume (million)	2.976	1.324	4.234	20.3	0.0032
Realised Volatility	0.59	0.303	0.835	5.528	0
Bid-ask Spd (bps)	18.13	16.99	9.253	55.99	3.565
Effective Spd (bps)	17.15	7.263	36.79	596.7	0.0604
HighLow (bps)	70.4	50.19	65.75	369.4	0
Depth (millon)	21	11.3	24.1	106	0.2165
Trade Size (shares)	3028	2748	1694	7760	336.2

Notes: The sample consists of 564 connected stocks and 454 non-connected stocks over the period from October to December 2014 (61 trading days). A firm-interval pair is dropped from the sample if there are trading suspensions on either the buy or sell side of that stock during the interval. The opening and closing auctions are excluded from the sample. Panel A reports interval average statistics for turnover, trading volume and realised volatility. Market capitalization is as of the end of December 2014. The table reports market capitalization at the end of 2014, and interval average statistics for turnover, trading volume, realised volatility, market quality and activity measures. Market depth and bid-ask spread are time-weighted averages for each firm during each interval. The effective spread is defined as twice the absolute value of the difference between the transaction price and the quote midpoint, and the average is value-weighted. The short-term volatility (HighLow) is defined as the highest midquote in an interval minus the lowest midquote in the same interval, divided by the midpoint between the high and the low, expressed in basis points. Trade size is the equally weighted average shares per transaction in each interval.

stocks with control variables are reported in Table 8.

The analysis of the bid-ask spreads and the effective spreads both show significant reduction after the SHHKConnect policy: the coefficients on the policy time dummy for both bid-ask spreads and effective spreads are negative and significant. The bid-ask spreads dropped by 2.45 bps and the effective spreads dropped by 1.71 bps for the cross-listed stocks after SHHKConnect. This result shows that cross-market competition plays a more important role than arbitrage. The benefits of increased competition outweigh the costs of cross-market arbitrage, because we observe narrower bid-ask spreads and effective spreads for cross-listed stocks after the policy. Even though the short-term volatility increased by about 5 bps after the policy, we still observe narrower effective spreads. Market depth has decreased slightly after the connect program by about 8.24%, indicating depth migration from the SSE to the SEHK.

In terms of market activity measures, the increase in the number of transactions per interval and average turnover per trade are significant, but small in magnitude, which shows an increase in trading activity for cross-listed stocks after the policy. We also observe a small but significant drop in trade size after the policy. Given the small drop in trade size, it is likely that the increase in average turnover per trade is related to the increase in stock prices during this sample period.

Overall, the market liquidity has improved after the SHHKConnect program for cross-listed stocks. The drop in effective spreads decreases the trading costs paid by investors. Competition for order flow between the two connected markets drives down both the quoted spreads and the effective spreads.

6 Further tests

6.1 Placebo test

In order to rule out the explanation that unobserved time-variant differences between connected and non-connected stocks drive the pattern of market quality measures, we implemented a placebo test. Specifically, we used the data before the SHHKConnect to conduct the test. We considered the pseudo announcement date to be 16 business days before the policy date, which is about three weeks before the SHHKConnect date, and repeat the

Table 8: Difference-in-differences regression results for cross-listed stocks

	Spread	EffSprd	HighLow	Depth	Transactions	TradeSize	Turnover
Policy	-2.453***	-1.712***	4.998***	-0.0824***	0.0477***	-0.215***	0.0869***
	(-32.55)	(-5.84)	(16.36)	(-9.55)	(17.24)	(-40.57)	(26.68)
RV	4.022***	4.003***	48.39***	-0.00916	0.110***	0.104***	-0.00975***
	(82.35)	(12.18)	(130.27)	(-1.61)	(56.46)	(32.3)	(-4.62)
lnTurnOver	-1.883***	-2.819***	17.00***	0.341***	0.742***	0.130***	
	(-49.66)	(-9.56)	(89.57)	(78.36)	(485.97)	(48.52)	
lnMarketCap	-0.564***	-3.036***	-9.578***	0.118***	-0.00890***	-0.118***	0.155***
	(-11.52)	(-11.77)	(-37.76)	(18.6)	(-3.84)	(-30.40)	(58.49)
lnVolume							0.0659***
							(41.49)
Constant	59.82***	127.8***	-24.75***	8.262***	-5.575***	8.385***	5.456***
	(62.39)	(23.94)	(-5.03)	(69.87)	(-116.49)	(117.29)	(95.1)
N	62587	62587	62587	62587	62587	62587	62587
R^2	0.134	0.0166	0.662	0.159	0.894	0.0954	0.139

This table reports the results of 10-minute interval panel difference-in-differences estimation for the cross-listed stocks on both the SSE and the SEHK. The dependent variables are bid-ask spreads (Spread), effective spreads (EffSprd), short-term volatility (HighLow), market depth (Depth), number of transactions in each interval (Transactions), average trade size (TradeSize), and average turnover per transaction (Turnover) for the connected and non-connected sample. The indicator variable picks out the period after the SHHK-Connect policy. The three control variables, realised volatility (RV), turnover (lnTurnOver), and market cap (lnMarketCap) are defined in Section 4.3. When turnover is the dependent variable, we used volume (lnVolume) as the control variable. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

difference-in-differences analysis. If there are certain unobserved time-variant factors other than the connect program that drive the relation we document, we would expect to observe similar relations in the pseudo dates as well.

The results of the placebo test are reported in Table 9. We find that the coefficients on the cross term become insignificant for effective spread and short-term volatility, which indicates that the connected stocks and matched non-connected stocks have indistinguishable changes in effective spread and short term volatility around the pseudo announcement date. However, the coefficients on the cross term for bid-ask spread and market depth are still significant. This suggests that the connected stocks and the matched non-connected stocks have unobserved time-variant differences in bid-ask spread and market depth before and after the pseudo announcement date. Because we are most interested in the influence of the policy on effective spreads, the placebo test reassures the observed increase in effective

Table 9: Placebo test results

	Spread	EffSprd	HighLow	Depth	Transactions	TradeSize	Turnover
Connected	-0.836***	-0.0163	-1.028***	0.0506***	0.157***	-0.0191***	-0.244***
	(-19.83)	(-0.04)	(-6.48)	(9.64)	(59.85)	(-5.74)	(-71.62)
Policy	-0.463***	0.185	-2.935***	0.0432***	0.0225***	0.00838***	0.0201***
	(-11.68)	(0.46)	(-19.68)	(8.77)	(9.12)	(2.68)	(6.27)
Connected_policy	0.614***	0.253	0.219	0.0254***	-0.00195	0.0233***	0.0196***
	(11.1)	(0.45)	(1.05)	(3.69)	(-0.57)	(5.34)	(4.38)
RV	7.854***	6.109***	55.10***	-0.0382***	0.194***	0.123***	0.0138***
	(265.32)	(20.47)	(494.41)	(-10.37)	(105.17)	(52.52)	(5.65)
lnTurnOver	-2.963***	-18.29***	10.98***	0.185***	0.627***	0.0900***	
	(-230.94)	(-141.40)	(227.33)	(116.18)	(786.39)	(88.86)	
lnMarketCap	1.368***	-6.755***	-9.881***	0.331***	-0.0183***	-0.0371***	0.312***
	(72.54)	(-35.52)	(-139.13)	(141.06)	(-15.64)	(-24.94)	(236.76)
Constant	29.08***	450.1***	75.78***	5.387***	-4.002***	6.946***	1.779***
	(105.04)	(161.27)	(72.71)	(156.43)	(-232.47)	(317.99)	(83.01)
N	274054	274054	274054	274054	274054	274054	274054
R^2	0.293	0.215	0.632	0.39	0.866	0.0728	0.411

Notes: This table reports the placebo test results of 10-minute interval panel difference-in-differences estimation with control variables. The dependent variables are bid-ask spreads (Spread), effective spreads (EffSprd), short-term volatility (HighLow), market depth (Depth), number of transactions in each interval (Transactions), average trade size (TradeSize), and average turnover per transaction (Turnover) for the connected and non-connected sample. Two indicator variables pick out connected sample stocks and the period after SHHKConnect policy respectively, and a further indicator variable interacts with the previous two. The three control variables, realised volatility (RV), turnover (lnTurnOver) and market cap (lnMarketCap) are defined in Section 4.3. The t-statistics are reported in parentheses. Statistical significance at the 10%, 5%, and 1% levels are denoted by one, two, and three asterisks, respectively.

spreads is indeed caused by the change in policy.

6.2 Parallel trends assumption test

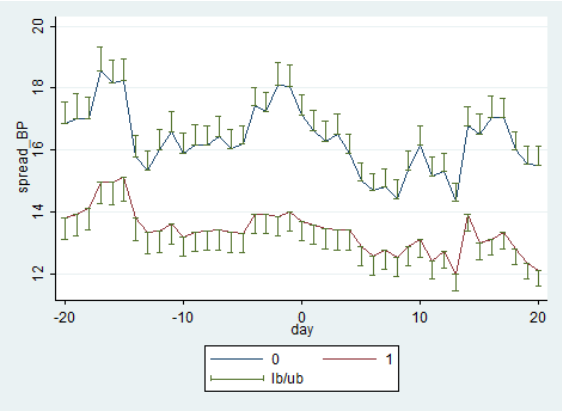
In a difference-in-differences model, the parallel trends assumption requires that in the absence of treatment, the difference between the 'treatment' and 'control' group is constant over time. Since the parallel trends assumption is critical to ensure validity of difference-in-differences models, we conduct a parallel trend test for all the dependent variables in the model to verify that the assumption is not violated.

We use the data for the connected and non-connected stocks over a 40-day period centered on the day the SHHKConnect was launched.[5] The results are reported in Figure 3.

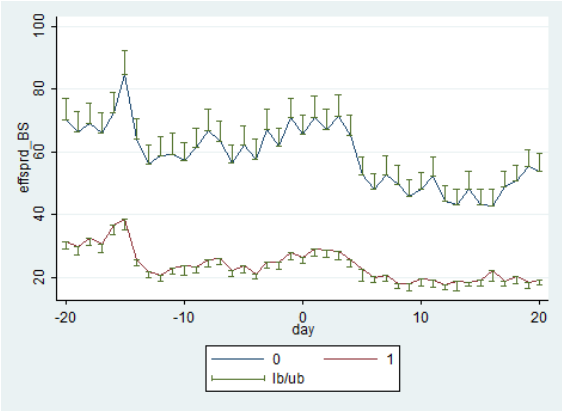
[5]The analysis for parallel trends over the whole sample period is available upon request. We only included the 40-day period results because the figures look clearer with less data observations and the results are very robust to the length of the time period chosen.

Figure 3: Parallel trends for the full sample

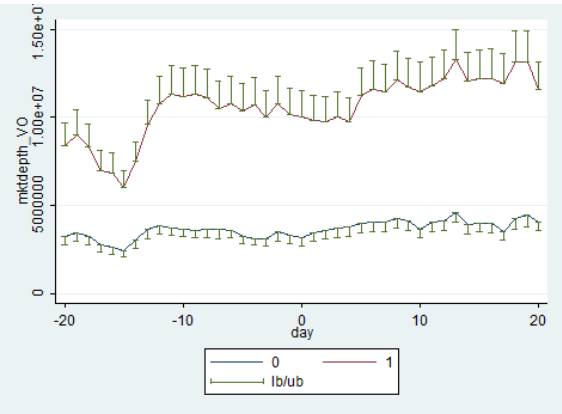
(a) Bid-ask Spread



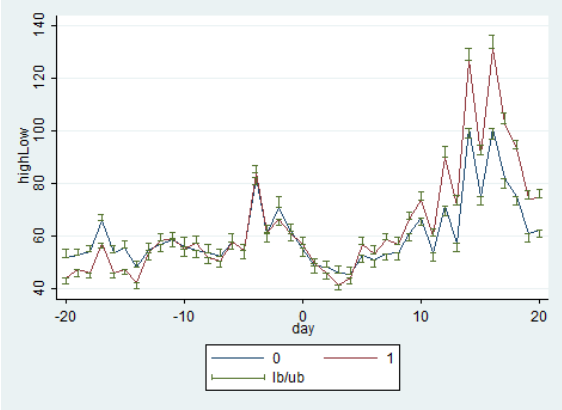
(b) Effective spread



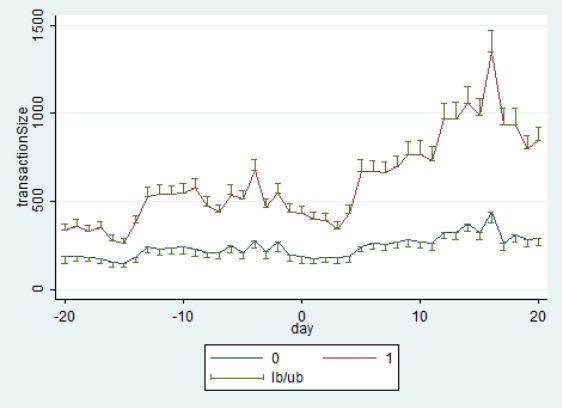
(c) Market depth



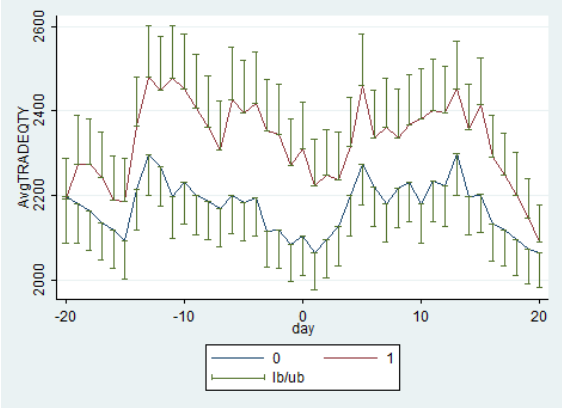
(d) Short-term volatility



(e) Number of transactions per interval



(f) Trade size



(g) Average turnover per trade

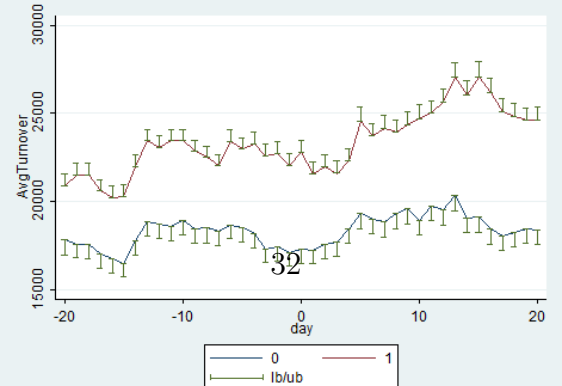


Figure 3 depicts four market quality measures and three market activity measures for the treated and control groups 20 days before the SHHKConnect and 20 days after the program. The vertical lines from each node reflect two standard errors of the mean values. Ideally, we want to observe the two lines representing the variables of interest for the treatment group and the control group trend closely in parallel in the 20 days leading up to the connect program. After SHHKConnect, the two lines start to converge and intersect, indicating a different change in the variables for the treatment group. Among all the seven variables we test, bid-ask spread, effective spread, short-term volatility, trade size, and turnover exhibit an obvious change in trend after the connect program. The number of transactions per interval seems to have a delayed change in trend about 5 days after the connect program. We do not observe a significant change in trend for market depth during the 40-day testing period.

7 Conclusion

This paper investigates the effect of capital market openness on market quality by studying a large market liberalization event in China: the Shanghai-Hong Kong Stock Connect (SHHKConnect). Using a difference-in-differences approach and exploiting the cross-sectional variation in market openness generated by SHHKConnect, we document that connected stocks in the Shanghai Stock Exchange experience significantly lower quoted spread, higher market depth, higher effective spread and higher short-term volatility compared to non-connected stocks with similar characteristics after the connect program.

To the best of our knowledge, our paper is the first to show that improved displayed liquidity does not necessarily lead to lower effective spreads because a lot of the limit orders are cancelled before being executed (fleeting orders). We empirically examine two possible mechanisms corresponding to market openness: competition channel and adverse selection channel, and find evidence for both. Improved displayed liquidity, as shown by lower bid-ask spread and higher market depth, is consistent with our competition hypothesis. Higher short-term volatility is evidence for more active cross-market arbitrage that increases adverse selection. We further examined the net effect of the two mechanisms by showing how market openness influences the investors' trading costs as measured by effective spreads. Our em-

pirical results indicate that the trading costs for Chinese domestic investors indeed go up as a result of market liberalization. More sophisticated foreign investors who are active in cross market arbitrage activities impose higher adverse selection costs for the retail-dominated Chinese investors. Although opening the market brings in more competition, the adverse selection channel dominates in China's stock market.

As the opening up of China's markets continues, academics and government regulators should recognize that market liberalization may lead to higher trading costs for Chinese investors. Our results suggest that it is reasonable for domestic investors to be concerned that the promotion of market openness comes with higher adverse selection costs for them. SHHKConnect paves the way for the integration of China's equity market with the global community, and removes barriers to entry. However, whether it allows Chinese domestic and foreign investors to compete on a level playing field is still an open research question. Further research in this area is warranted.

Appendix A. List of Model Variables

This appendix summarizes key variables used in the paper.

Variable	Description
$BidAskSpd$ or $Spread$	Best ask price minus best bid price
$EffSpd$	Difference between the transaction price and the quote midpoint
$HighLow$	The highest midquote minus the lowest midquote in an interval
$Depth$	Number of shares in the book up to 10 cents from the best quotes
$TradeSize$	Number of shares per trade
$Transactions$	Number of transactions per 10-minute interval
$Turnover$	Average turnover per transaction
$MarketCap$	Market capitalization as of the end of Dec 2014
$Volume$	Average trading volume per interval
$RealisedVolatility$	Sum of squared returns in each interval
$Connected$	Equals one for connected stocks and zero for non-connected stocks
$Policy$	Equals one if after the SHHKConnect program, otherwise zero
$Connected \times Policy$	A cross term dummy variable of the previous two dummies
α_i	Stock specific fixed effects
λ_t	Time fixed effects

References

Angel, J. J., L. E. Harris, and C. S. Spatt (2011). Equity trading in the 21st century. *Quarterly Journal of Finance 1*(01), 1–53.

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

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

Bai, Y. and D. Y. P. Chow (2017). Shanghai-hong kong stock connect: An analysis of chinese partial stock market liberalization impact on the local and foreign markets. *Journal of International Financial Markets, Institutions and Money 50*, 182–203.

Baldauf, M. and J. Mollner (2017). Trading in fragmented markets. *Available at SSRN 2782692*.

Bekaert, G. and C. R. Harvey (2000). Foreign speculators and emerging equity markets. *Journal of Finance 55*(2), 565–613.

Bekaert, G., C. R. Harvey, and C. Lundblad (2011). Financial openness and productivity. *World Development 39*(1), 1–19.

Biais, B., T. Foucault, and S. Moinas (2015). Equilibrium fast trading. *Journal of Financial Economics 116*(2), 292–313.

Brogaard, J. and C. Garriott (2014). High-frequency trading competition. *Journal of Financial and Quantitative Analysis*, 1–66.

Budish, E., P. Cramton, and J. Shim (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics 130*(4), 1547–1621.

Burdekin, R. C. and P. L. Siklos (2018). Quantifying the impact of the november 2014 shanghai-hong kong stock connect. *International Review of Economics and Finance*.

Carpenter, J. N., F. Lu, and R. F. Whitelaw (2017). The real value of china's stock market. *Work. Pap., Stern Sch. Bus., NY Univ. Article Locations: Article Location Article Location Article Location Article Location Article Location Article Location Article Location Article Location Article Location*.

Chan, M. K. and S. Kwok (2017). Risk-sharing, market imperfections, asset prices: Evidence from china's stock market liberalization. *Journal of Banking and Finance 84*, 166–187.

Colliard, J.-E. and T. Foucault (2012). Trading fees and efficiency in limit order markets. *Review of Financial Studies 25*(11), 3389–3421.

Doidge, C., G. A. Karolyi, and R. M. Stulz (2004). Why are foreign firms listed in the us worth more? *Journal of Financial Economics 71*(2), 205–238.

Dutta, P. K. and A. Madhavan (1997). Competition and collusion in dealer markets. *Journal of Finance 52*(1), 245–276.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market1. *Journal of Financial Markets 2*(2), 99–134.

Foucault, T., R. Kozhan, and W. W. Tham (2017). Toxic arbitrage. *Review of Financial Studies 30*(4), 1053–1094.

Friederich, S. and R. Payne (2015). Order-to-trade ratios and market liquidity. *Journal of Banking and Finance 50*, 214–223.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *Journal of Finance 50*(4), 1175–1199.

Hasbrouck, J. (2018). High-frequency quoting: short-term volatility in bids and offers. *Journal of Financial and Quantitative Analysis 53*(2), 613–641.

Hasbrouck, J. and G. Saar (2009). Technology and liquidity provision: The blurring of traditional definitions. *Journal of Financial Markets 12*(2), 143–172.

Hasbrouck, J. and G. Saar (2013). Low-latency trading. *Journal of Financial Markets 16*(4), 646–679.

Henry, P. B. (2000). Stock market liberalization, economic reform, and emerging market equity prices. *Journal of Finance 55*(2), 529–564.

Ho, D. E., K. Imai, G. King, and E. A. Stuart (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis 15*(3), 199–236.

Ho, T. S. and H. R. Stoll (1983). The dynamics of dealer markets under competition. *Journal of Finance 38*(4), 1053–1074.

Hui, E. C. and K. K. K. Chan (2018). Does the shanghai–hong kong stock connect significantly affect the ah premium of the stocks? *Physica A: Statistical Mechanics and its Applications 492*, 207–214.

Huo, R. and A. D. Ahmed (2017). Return and volatility spillovers effects: Evaluating the impact of shanghai-hong kong stock connect. *Economic Modelling 61*, 260–272.

Jørgensen, K., J. Skjeltorp, and B. A. Ødegaard (2017). Throttling hyperactive robots–order-to-trade ratios at the oslo stock exchange. *Journal of Financial Markets*.

Larrain, M. and S. Stumpner (2017). Capital account liberalization and aggregate productivity: The role of firm capital allocation. *Journal of Finance 72*(4), 1825–1858.

Lin, W. (2017). Modeling volatility linkages between shanghai and hong kong stock markets before and after the connect program. *Economic Modelling 67*, 346–354.

Liu, C., S. Wang, and K. J. Wei (2016). Demand shock, speculative beta, and asset prices: Evidence from the shanghai hong kong stock connect program. *Working paper*.

Malinova, K., A. Park, and R. Riordan (2016). Taxing high frequency market making: Who pays the bill? *Working paper*.

Menkveld, A. J. (2018). High-frequency trading as viewed through an electron microscope. *Financial Analysts Journal 74*(2), 24–31.

Ng, A. (2000). Volatility spillover effects from japan and the us to the pacific–basin. *Journal of International Money and Finance 19*(2), 207–233.

Pagnotta, E. S. and T. Philippon (2018). Competing on speed. *Econometrica 86*(3), 1067–1115.

Sohn, S. and N. Jiang (2016). Stock market liberalization and price discovery: Evidence from the shanghai-hong kong stock connect. *Available at SSRN 2850967*.

Stiglitz, J. E. (2010). Risk and global economic architecture: Why full financial integration may be undesirable. *American Economic Review 100*(2), 388–92.

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

Xu, K. (2018). Who makes the market during stressed periods? hfts vs. dealers. *Available at SSRN 3045506*.

Zhang, F. (2010). High-frequency trading, stock volatility, and price discovery. *Available at SSRN 1691679*.

Zhang, Q. and S. A. Jaffry (2015). High frequency volatility spillover effect based on the shanghai-hong kong stock connect program. *Investment Management and Financial Innovations 12*(1), 8–15.