The Impact of Latency Sensitive Trading on High Frequency Arbitrage Opportunities

Alex Frino

Macquarie University Graduate School of Management

Vito Mollica

Macquarie University Graduate School of Management

Robert I. Webb

University of Virginia

Shunquan Zhang

Macquarie University Graduate School of Management

May 31, 2015

Abstract

This study examines the *duration*, *frequency* and *profitability* of potential arbitrage profit opportunities between the Australian Securities Exchange (ASX) Share Price Index futures contract and an exchange-traded fund (ETF) written on the S&P/ASX200 constituent securities. Not surprisingly, we find the frequency and profitability of potential arbitrage opportunities are greater during volatile and high turnover periods - other things equal. We examine the increased competition in high frequency trading by identifying the number of 'cabinets' co-located in the ASX's liquidity center. With increased HFT connections, we observe increasing value, frequency and duration of index arbitrage profit opportunities. Our results are robust to the inclusion of transaction costs. We conclude that the activity of disruptive HFT outweighs the activity of index arbitrage HFTs.

The Impact of Latency Sensitive Trading on High Frequency Arbitrage Opportunities

1. Introduction

Classical economic theory suggests that excess returns should be competed away as new participants enter the market. This is especially true for the profits from riskless arbitrage. Yet, there is conflicting evidence in the financial economic literature over whether high frequency trading (HFT) profits, in general, (Baron et al [2012]) and arbitrage profits, in particular (Budish et al [2013] and Chaboud et al [2013]), decline as high frequency or other algorithmic trading increases. There are important public policy implications for market microstructure and the social value of investments by HFT firms in being faster if arbitrage profit opportunities persist (in the absence of limits to arbitrage).

There are several different strategies that high frequency and other latency sensitive traders engage in. These include: index arbitrage; spread arbitrage/market making; and correlated arbitrage among others. This study focuses on only one-- index arbitrage. Specifically, it examines whether the *duration*, *frequency* and *profitability* of potential arbitrage opportunities between the Australian Securities Exchange (ASX) Share Price Index (*SPI*) futures contract and the exchange traded fund (*ETF*), *STW*[1], have changed as the number of HFT firms (or intensity of HFT activity) has increased, since the ASX's introduction of co-location services in February 2012. In addition, we use estimated potential arbitrage profits and compare them to the cost of being co-located to determine the value of minimum latency.

We have two principal reasons for examining ASX data. First, we are able to examine a finer time interval than past studies, with data time-stamped at the microsecond level (i.e. one millionth of a second). Second, we have information on the growth of HFTs market involvement over our sample period. More specifically, we know the number of colocated cabinets reported by the ASX in its minimum latency liquidity centre and use this measure as a proxy for HFT competition.

Our principal findings are as follows: First, consistent with Budish et al [2013], the *frequency* of potential arbitrage opportunities is greater during volatile periods, other things equal. Second, increased HFT in the market changes the speed of convergence between the paired instruments, suggesting greater competition amongst HFTs in demanding or providing liquidity. Third, the average daily arbitrage profit, frequency and duration increase, as more HFTs connect to the market, proxied by occupied collocated cabinets. HFT firms appear to compete fiercely against each other, resulting in larger and

1 SPDR® S&P®/ASX 200 Fund

more frequent price discrepancies. Our findings are consistent with Kozhan & Tham [2012] who demonstrate that HFTs suffer execution uncertainty from negative externalities inflicted by trading against each other.

2. Review of the Literature

Baron, Brogaard, and Kirilenko [2012] examine the profitability of HFTs in the (all electronically traded) e-mini S&P 500 stock index futures market during the entire month of August 2010. Using a comprehensive data set identifying the trades of 31 HFT firms, they report all 31 HFT firms were profitable during the month. The firms collectively earned $29 million during the month all while assuming very little risk. Indeed, the average Sharpe ratio was 9.2. The most profitable HFT firms were the most aggressive who were primarily liquidity takers rather than liquidity providers. Baron et al [2012] argue that their results provide evidence against market efficiency. They state: "… the magnitude of their profits suggests that HFTs still earn significant excess returns even after accounting for their costs. The magnitude of their profits brings into question the efficiency of markets at high-frequency time scales."

Baron et al [2012] argue: "How can HFTs maintain consistently high profits without having their profits driven down by competition? Note that while Passive HFTs have high Sharpe ratios, their magnitude of profits are relatively small, suggesting that competition for liquidity provision may be driving down profits. In contrast, aggressive HFTs, who earn the highest profits, seem to be unaffected by competitive forces. It may be that the profitability of aggressive HFTs depends on their relative speed: in a winner-takes-all market, profits accrue almost entirely to the fastest. The speed and technological sophistication needed to compete with the most profitable firms may represent a barrier to entry in the HFT market."

The focus on the continuing investments in speed by HFT firms as a potential explanation for why profits remain relatively stable extends to Budish et al [2013]. Essentially, Budish et al [2013] argues that correlation between related securities breaks down in continuous markets over very short periods of time and "creates purely technical arbitrage [profit] opportunities, available to whomever is fastest, which, in turn, create [a socially wasteful] arms race to exploit these arbitrage opportunities." Budish et al [2013] estimates the annual technical arbitrage profit opportunities across markets are "in the billions [of dollars]."[2]

Budish et al [2013] argues that the duration of one such technical arbitrage opportunity (between the S&P 500 e-mini futures contract and an exchange traded fund for the S&P

[2] The estimate by Budish et al [2013] of billions of dollars of technical arbitrage profit opportunities conflicts sharply with industry estimates of the total profits of high frequency trading firms in the U.S. equity market Tabb [2014] estimates that profits among equity HFT firms in the U.S.A. have declined sharply from $7.2 billion to $1.3 billion. If this industry estimate is correct then either HFT firms are leaving substantial profits on the table or Budish et al [2013] are incorrect in their estimate of the size of such profit opportunities. Given that access to real-time information is a necessary condition to conduct low-latency trading, the question naturally arises as to whether exchanges are pricing their colocation services correctly if Budish et al [2013] are correct that there are billions of dollars of purely technical arbitrage profit opportunities. See Webb [2003] for a discussion of the history of exchange ownership of real-time price and volume data.

500 index), has declined but the profitability of such "arbitrage opportunities is remarkably constant" although the frequency varies directly with market volatility. Simply put, competition among HFTs does not bid away the potential arbitrage profit opportunities as HFTs compete on investments in being faster vis-a-vis other traders rather than providing better prices. Budish et al [2013] use the finding of no change in the dollar profitability of potential arbitrage profit opportunities to advocate replacing continuous trading with a series of high frequency batch auctions spaced as little as 100 milliseconds, or as much as 1 second, apart to reduce the *sniping costs* in the bid offer spread that result from the race to be fastest.

Budish et al [2013] compare the prices of the e-mini S&P 500 stock index futures contract which is traded on the Chicago Mercantile Exchange (CME) with prices of the SPDR ("Spider") S&P 500 stock index ETF which is traded on multiple venues including the New York Stock Exchange (NYSE) under the ticker symbol SPY. They note that they use direct feeds from the CME and the NYSE in conducting their empirical analyses. The question naturally arises as to whether the data are truly synchronized especially since Budish et al limit their empirical analysis to only SPDR prices reported on the NYSE.[3] Simply stated, if the two time series are not synchronized "observed" arbitrage profit opportunities may be more apparent than real. And, the presumption that there is a large amount of HFT index arbitrage activity may be incorrect.[4] There may also be limits to arbitrage that impede HFTs from exploiting apparent arbitrage opportunities as Kozhan and Tham [2012] point out.

In contrast, Chaboud et al [2013] report evidence of a substantial decline in the frequency (and implicitly the amount) of triangular arbitrage profit opportunities in the foreign exchange market. Indeed, Chaboud et al [2013] report many instances where no triangular arbitrage profit opportunities exist. Chaboud et al [2013] attribute the decline in arbitrage profit opportunities to the growth of algorithmic trading. They argue that by reducing the time it takes for price discovery algorithmic trading increases informational efficiency. To be sure, their use of "minute-by-minute" data may obscure a decline in the *duration* of potential arbitrage profit opportunities even though the overall *amount* of

[3] Miller, Muthuswamy and Whaley [1994] examine index arbitrage in the S&P 500 stock index futures market and point out some of the dangers of inferring arbitrage when data are not synchronized due to infrequent trading. They illustrate the problem of infrequent trading with the following example: "Basis reversions of this kind, unrelated to arbitrage, have long been recognized when they occur at the openings on days with a heavy imbalance of orders like Monday, October 19, 1987. The futures market opened that day down seven percent, ... The reported index did not fall immediately, however, because it is based on the last transaction price of each component stock, and some large capitalization stocks in the index, including IBM, did not trade at the regular opening. The index was thus reporting mainly the long-since obsolete prices of Friday's close, not the prices actually achievable at Monday's opening. As each stock in turn opened down by seven percent, the reported index level moved closer to the futures price. But the process was slow. Ninety minutes passed before the reported basis returned to its equilibrium value."

[4] Miller et al [1994] note: "But, if the predictability of basis changes is mainly a statistical illusion, as we have argued, why do we see so much index arbitrage on the NYSE? … The answer, we have shown, is that we don't really see all that much of it. Formal index arbitrage during our sample period accounts for only about four percent of NYSE volume. And such formal arbitrage as we do see appears to serve mainly to counteract the additional drag in index adjustment induced by the very special set of rules that the NYSE imposes to create the impression of continuity in the path of prices. In a pure dealer market such as the spot market in Treasury bonds, no actual arbitrage transactions are needed to keep the spot and futures prices in line. The dealers, upon observing a jump in futures price, would simply mark their own stale quotes up or down to match. An exploitable gap does not emerge, because, effectively, the spot dealers "price off the futures" and hence eliminate any profit opportunity directly."

arbitrage profit opportunities is unchanged. That is, declines in the duration of arbitrage opportunities may result in an apparent absence of arbitrage profit opportunities at the minute-by-minute level that are apparent at the millisecond level.

3. Data and Methodology

The data used in this study are sourced from Thomson Reuters Tick History (*TRTH*), managed and distributed by *SIRCA*. We sample tick level data, time stamped to the nearest microsecond, containing price and volume information of each best bid/ask update over the period of March 1, 2012 to January 31, 2014. Two closely related instruments are examined: 1) *STW*, an Exchange Traded Fund (*ETF*) tracking ASX/S&P 200 index, and 2) *SPI*, an index future contract written over the identical index.

STW is managed by *State Street Global Advisors* and available for trading on Australian equity platforms, ASX and Chi-X.[5] Chi-X commenced trading ETFs and other equities in October 2011. In the subsequent analysis, we assume HFTs have direct market access to both venues and thus construct a consolidated tape from the two markets to provide the best representation of the national limit order book.[6]

On any given day, multiple *SPI* future contracts are available in the market with different expiry dates. We sample only the most actively traded contract on each trading day. *SPI* trades solely on the ASX's derivative platform, ASX Trade24, over two sessions, a day session (9:50am – 4:30pm) and an overnight session (5:10pm – 7:00am). We only examine the day session which overlaps with the corresponding ETF trading hours. We also exclude trades during open and close auctions on the equity platform. This results in a daily sample period extending 10:10am to 4:00pm, during which time both instruments are available for continuous trading.

a. Measuring High Frequency Trading

To identify the level of HFT competition in the Australian market, we utilize publicly available information stipulating the usage of the ASX's colocation facility. ASX introduced its colocation facility, the Australian Liquidity Centre (ALC), for equities products (including ETFs) on February 6, 2012, and futures and options products on February 20, 2012. ALC hosts 'cabinets' for its customers with minimum latency access to the ASX matching engine and order book, a prerequisite for any HFT operation.[7] The number of cabinets rented is reported by ASX in its yearly and half-yearly reports. There were 76 cabinets occupied as at the end of June 2012, the first fiscal year of co-location; it increased to 111 by December 2012, 117 by June 2013 and 133 by December 2013. At the commencement of the ALC colocation facility, there were in total 20 occupied

[5] *STW* does not have any exchange traded options or future contracts.

[6] Australia does not operate under a Reg NMS structure as in the US. Our results are robust to the exclusion of any Chi-X data, which typically represents 10 percent of daily traded volume.

[7] It can be argued that some market participants may strategically rent a large number of colocation cabinets, without truly occupying the space. However ALC has a capacity of 300 colocation cabinets, almost twice the size of current utilised capacity, rendering such strategy pointless.

cabinets, which were moved from ASX's previous data center[8]. These dates form event windows for the subsequent analysis.

b. HFT Arbitrage

In theory, *SPI* and *STW* should track one another perfectly, since they reference the same underlying securities. Discrepancies in price levels between the two arise due to cost of carry, dividends, ETF tracking error in the basket of underlying securities, and contract specifications.[9] Panel A, Figure 1 shows the price path of the pair over the entire sample period. *SPI*'s price is consistently above *STW* and daily price pattern depicts a near perfect parallel shift. This parallel breaks down at finer trading intervals. Panel B, Figure 1 displays the price sequence from 2pm to 3pm, on a randomly chosen trading day. Prices still appear to track each other closely, suggesting high correlation between intraday prices of the two. However, the unsynchronized movements of the two instruments at finer snapshots highlight the existence of temporary price discrepancies and potential arbitrage opportunities.

[INSERT FIGURE 1]

Given the high correlation depicted in Figure 1, any transitory price discrepancies between the two instruments should be quickly resolved by market participants' updating their orders. Discrepancies which are not immediately adjusted can be captured (traded) by other participants in the market. For instance, if the *SPI* experiences a sudden increase and *STW* remains the same, one can short *SPI* and long *STW*, and liquidate the position following the price reversion. Since the correlation converges rather rapidly, the time it takes to eliminate such a discrepancy is reasonably short. This style of statistical arbitrage can be automated by HFTs who invest in minimum latency technology to be the first to take the advantage of mispricing, and is the focus of our study.[10] We are able to identify these potential arbitrage opportunities and measure the frequency of arbitrage, potential profit and duration of each occurrence.

The first step to identify an arbitrage is to quantify the natural price difference between *SPI* and *STW*. Following Budish et al [2013], we start by defining the immediate spread as the price difference at any time point *t*, being,

$$S_t = P_t^{SPI} - 100P_t^{STW},$$

[8] ASX established a data centre, in Bondi, Sydney, with a maximum capacity of 20 hosts, which was fully occupied prior to the development of the ALC. Occupants of the old data centre were relocated to the new co-location cabinets after ALC became operational. See http://www.itnews.com.au/News/289358,new-asx-data-centre-goes-live.aspx.

[9] There are two major differences between *SPI* and *STW*. While relative minimum tick size (minimum tick divided by price) for *SPI* and *STW* are very close, the dollar value of the *SPI* is substantially larger than the dollar value of the *STW*. A one tick movement in *SPI*, one index point, results in a change of $25(AUD) per contract, while the equivalent one tick change in *STW*, represents only a 1 cent change per unit of *STW*. This means that one contract of *SPI* is equivalent to 2500 *STW* units. Lastly, the expiration of *SPI* contracts at the end of each quarter requires traders to roll over to the next contract if they wish to continue holding their position. *STW*, on the other hand, pays out accumulated dividends and is required to rebalance its portfolio, by the issuing company, to re-weight changes in ASX200 index basket at the end of each quarter.

[10] HFTs typically avail themselves of three strategies: (1) single stock spread capture (2) correlated security spread capture and (3) index arbitrage.

where P_t^{SPI}, P_t^{STW} are the mid-quote price levels of the two instruments, and *100* reflects the quoted price differences in *SPI* and *STW*. A one-point increment in the ASX 200 index is equivalent to one index point increasing in SPI futures and \$0.01 increasing in the quoted price of STW.

The immediate spread (S_t) is however not sufficient to be the basis of arbitrage as its updating frequency, while timely and relevant, is too volatile and noisy to be used as a reliable signal. By accounting for the average spread over a period of time, prior to the current best bid and ask update, one can best identify more consistently realizable arbitrage opportunities. This chosen time period should be at a level that is long enough for the two prices to converge (a stable average spread), but also short enough for the information it contains to be relevant to the current update. In this paper, we determine this time parameter, r, the time it takes for the returns of two instruments to reach a correlation of 0.90, during the 20 trading days prior to the current date[11]. More formally, on a particular trading day, i, then the time gap r_i is determined by solving the equation $Corr(r_i) = 0.90$, accurate to the millisecond level, using data from the prior 20-trading-days.[12]

The basis of arbitrage opportunities at each price update is then valued as the time weighted spread, $\overline{S_t}$, during the previous r seconds as,

$$\overline{S_t} = \frac{1}{r}\sum_{i=1}^{n_r} S_{t-t_i} \cdot \{t_i - Max(t_{i-1}, t_{n_r})\},$$

where: $\{n_r | t - t_{n_r - 1} < r \leq t - t_{n_r}\}$. $\{t_i - \text{Max}(t_{i-1}, t_{n_r})\}$ represents the time gap between each price update, with the exception of the first observation, since r almost never coincides precisely with the time of quote updates. n_r indicates the number of previous spread updates to be included. If a sudden change is observed in either market, which is then reverted back to its starting value within r seconds, the average spread $\overline{S_t}$ still remain unchanged.

The basis spread, $\overline{S_t}$, is then compared against the update of the best bid and ask at time *t*. We define the bid and ask of the arbitrage strategies as $S_t^{bid} = {}_{bid}P_t^{SPI} - 100\,{}_{ask}P_t^{STW}$ (short *SPI* & long *STW*), and $S_t^{ask} = {}_{ask}P_t^{SPI} - 100\,{}_{bid}P_t^{STW}$ (long *SPI* & short *STW*). By design, the relationship $S_t^{bid} < \overline{S_t} < S_t^{ask}$ holds when there is no price divergence. If, for instance, there is a sudden decline in *SPI* (or increase in *STW*) at time u, which is large

[11] Returns are valued based on the sampled price sequence. Generally, price sequence can be sampled as beginning, ending and average price over the interval. Throughout the paper, we adopt the time weighted average price approach to ensure the sampled prices being more representative of the individual time interval.

[12] There are a few trading days in our sample where convergence in returns is slow (e.g., days with *STW* trading halts). To ensure r_i is not influenced by erroneous data errors, we remove trading days, which require more than 2,000 seconds to reach a correlation of 0.99. This excludes less than 5% of the sample period. Secondly, as shown in Figure 1, the correlation does not act as a continuously increasing function against the time interval of the return, because it fluctuates. We use a binary search method to find a solution to $Corr(r_i) = 0.90$. Binary search takes approximately 21 iterations of function evaluations to find the required accuracy, while conventional methods would require a few hundred thousand such evaluations. A caveat with this approach is that it does not guarantee the smallest r_i.

enough so that $S_u^{bid} < S_u^{ask} < \overline{S_u}$, an arbitrage profit opportunity emerges, in the direction of S^{ask}. The length of such an opportunity ceases when the spread moves back to normal or reverts at some future time v. The difference $v - u$ is referred as the *duration per arbitrage*.

Once the spread reverts and the position is closed out, the potential profit from such a strategy at the starting time u, measured in index points per contract is $\overline{S_u} - S_u^{ask}$. The reverse is true when *SPI*/*STW* price suddenly rises/declines.[13] The profit of the arbitrage can be summarized as follow,

$$\pi = \begin{cases} \overline{S_u} - S_u^{ask}, & if\ \overline{S_u} > S_u^{ask} \quad \Rightarrow long\ SPI\ \&\ short\ STW \\ S_u^{bid} - \overline{S_u}, & if\ \overline{S_u} < S_u^{bid} \quad \Rightarrow short\ SPI\ \&\ long\ STW \end{cases}.$$

The change in spread S_t does not always converge back, due to a permanent shock in the *SPI*/*STW* price discrepancy. We adopt the approach by Budish et al [2013], if the arbitrage duration, $v - u$, is larger than r, then such arbitrages are deemed to be "bad arb".

Expected *Dollar profit per arbitrage* is then calculated by multiplying π with the existing tradable $Volume$ defined as,

$$Volume = \begin{cases} 25 \times Min\left(Vol_{SPI}^{ask}, \dfrac{Vol_{STW}^{bid}}{2{,}500}\right) & if\ \overline{S_u} > S_u^{ask} \\ 25 \times Min\left(Vol_{SPI}^{bid}, \dfrac{Vol_{STW}^{ask}}{2{,}500}\right) & if\ \overline{S_u} < S_u^{bid} \end{cases},$$

where: 2,500 adjusts the *STW* share to the same value as per *SPI* contract and 25 reflects the dollar value per index point. *Daily profit* is summed and the total *frequency of arbitrage* occurrences is counted at the end of each trading day. *Duration per arbitrage* is aggregated on a daily basis into *Daily Duration.*

c. Analysis

While the above process identifies potential arbitrages and mispricing in the market and can be applied over our event windows, we control for a number of market factors in order to investigate the relation between arbitrage opportunities (mispricing) and HFTs. Specifically, we control for the turnover in *SPI*, since *SPI* contracts dominate in both trading volume and the process of creating arbitrage opportunities. Consistent with Brailsford and Hodgson [1997], Cummings and Frino [2011] and Baron, et al [2013] we control for volatility in futures, measured by taking the log of the ratio between the daily highest and lowest mid-price in *SPI* contract.[14] Scholtus, van Dijk, and Frijns [2014] find algorithmic trading increases around macroeconomic news releases. To control for news

[13] Note that the calculation of S_t^{ask} and S_t^{bid} uses the information of price updates at time t, while $\overline{S_t}$ only contains price information prior to time t.

[14] We also included the ratio of volatility in the *SPI* and *STW*, the variable is insignificant and does not alter results.

days, we include an impact measure developed by Bloomberg that measures the relevance of macro-economic news from 0 to 100.[15] We also include the previous day S&P/ASX 200 index return to capture any effect of market price momentum on the amount of potential arbitrage profit opportunities in the market.

Similar to the existing research examining mispricing, we estimate the parameters of the following model using OLS regression:

$$Arb_t = Intercept + \beta_1 Colocation_t + \beta_2 \log(1/SPI\ Dollar\ Volume_t) + \beta_3 Volatility_{SPI,\,t} + \beta_4 NewsRelevence_t + \beta_5 LagReturn_t,$$

where: Arb_t is any of the three daily arbitrage variables defined earlier: *daily duration*, *daily arbitrage profit* and *arbitrage frequency* per day; $Colocation_i$ is the number of cabinets hosted during each event window and our proxy for HFT competition.[16]

Our analysis considers five event windows as follows: Window 1 is for the month of March 2012, Window 2 extends May 2012 to July 2012, Window 3 extends November 2012 to January 2013; Window 4 extends May 2013 to July 2013, and Window 5 extends November 2013 to January 2014. Although exact utilization days of individual ALC cabinets are not in the public domain, each event window (except window 1) represents one-month before and after the reporting date of ASX Financial Reports, so on average the occupied cabinets during each event window reflects the reported colocation figures.[17]

4. Results

We begin our analysis by reporting return correlations between the two instruments-SPI and ETF. Figure 2 depicts median daily correlation using mid-price returns sampled at various time intervals. Panel A presents returns sampled from 1 to 1000 milliseconds and Panel B shows the correlation of returns from 1 to 11 seconds with 10 millisecond intervals, followed by Panel C, 11 to 111 seconds measured over 100 millisecond increments. Over the sample period, March 2012 to January 2014, the median daily correlation between the *SPI* and *STW* starts at 0.154 at 1 millisecond and rises to 0.570 by 1 second, 0.765 by 11 seconds and 0.936 by 111 seconds.

[INSERT FIGURE 2]

Examining this behavior over our five event windows, we observe a more rapid convergence in returns as the number of collocated cabinets increases. Such behavior is particularly evident over finer time intervals (Panel A). However, this does not appear to hold for the May 2013 – July 2013 window. Compared to results in Budish et al., [2013],

[15] For days with more than one announcement, we sum the total relevance score to form the variable $NewRelevence_i$
[16] The regressions are Newey-West (1987) adjusted for Heteroskedasticity. Outliers of the three dependent variables, defined as two standard deviation away from the mean, are removed from the analysis.
[17] The regression analysis is also carried out under the assumption of a linear increase in cabinet utilization for all the trading days in the sample period (Mar 2012 – Jan 2014), and similar results are observed.

who report near zero median correlation for returns measured over 1 millisecond in the US, we observe a minimum median correlation of 0.154 in Australia. Beyond these fine intervals, the speed of convergence is much slower in Australia vis-à-vis the US. For example, Budish et al [2013] finds correlation to be approximately 0.90 by 11 seconds, while in Australia the correlation is only 0.765.

[INSERT TABLE 1]

Panel A, Table 1 reports descriptive statistics for *SPI* and *STW* over the sample period. The *SPI* futures market is more active and liquid than the *STW* ETF market. The time weighted bid-ask spread in *SPI* is 1.13 index points, 13 percent larger than the minimum tick. The bid-ask spread in *STW* is, on average, 2.24 cents, 124 percent larger than the minimum tick and nearly twice as large as that observed in the *SPI*[18]. The notional value of the depth at the best bid and ask prices for the *SPI* is on average $4.67 million, vis-à-vis a notional value of the depth in the *ETF* of $1.76 million. The *SPI* futures have, on average, a daily dollar turnover 200 times the size of the *STW* and is 23.21 times more frequently traded. The number of best bid and offer updates on the *STW* is typically around 12,840 messages per day, while it is 74,500 for the *SPI*. The difference between these two numbers impacts the speed of convergence in correlation between the pair of instruments. Daily volatility of the two instruments is relatively similar as expected, given they track the same underlying index. Panel B, Table 1 reports summary statistics in each of the five event windows, and the results are similar to those in Panel A. We do however note heightened bid-ask spreads and quote updates in Window 4.

Employing the methodology introduced in section 3, we quantify the arbitrage opportunities. Summary statistics are reported in Table 2.

[INSERT TABLE 2]

On average, our strategy identifies 124 potential arbitrage profit opportunities on a daily basis, which accumulates to a potential profit of $529 per trading day. On average, arbitrage opportunities last 606 seconds per day. *Convergence Time (*i.e., *r*) records the time used in each trading day to calculate the basis spreads between *SPI* and *STW*. It takes 75 seconds to reach a correlation of 0.90 over the previous twenty trading days, ranging from 16 seconds (1st percentile) to 165 seconds (99th percentile). Over 80% of the apparent arbitrage profit opportunities arise from quote price movements in the *SPI*. This is consistent with the greater trading activity in the *SPI* as compared to the *STW*. *% Good Arbs* shows the percentage of arbitrage opportunity that reverse before the daily convergence time, r. More than 98% of the recorded arbitrages are deemed to be "good", which demonstrates the close correlation of the two instruments and stability of the adopted arbitrage mechanism.

---

[18] During the entire sample period, *STW* is traded on both the Chi-X and ASX without exchange established, inter-market automated best bid and ask prices. This means that the real minimum spread of STW can be zero. This can happen when the best bid on Chi-X and best ask on ASX are temporarily at the same price level. "Market makers" would not exploit such an opportunity, since the mispricing does not cover the transaction cost. Such opportunities would only disappear when participants on one side of the spread realize the available liquidity in the other market and cancel or resubmit their orders. If the market was not fragmented, the same orders would result in a spread of one or possibly two ticks. The real spread in *STW*, traded in a single market, is thus even larger than reported.

[INSERT TABLE 3]

Table 3 summarizes arbitrage (*Profit, Duration and # Arbitrage per day*) and control variables, during each of the five event windows and over the entire sample period. The variables do not appear to exhibit a clear trend, e.g. the average duration is at 232.9 seconds at the start (March 2013) and increases to 489.6 (May – Jul 12) and 681.5 seconds (Nov 12 – Jan 13). The average arbitrage time then falls to 370.4 seconds by mid-2013 and increases to 749.9 seconds at the end of 2013. Additionally, we observe that the arbitrage opportunities of the pair *SPI*/*STW* are, on average, one six-hundredth of the daily profit figure of reported by Budish et al [2013], reflecting the larger trading size and liquidity of the futures and ETF contracts in the US.

Table 4 Panel A reports regression results. Our key variable of interest, *Colocation*, measures the impact of HFT competition on arbitrage profit opportunities. Firstly, we find an increase in co-located cabinets is associated with an increase in overall daily profit, duration and frequency of potential arbitrage opportunities between *SPI* and *STW*. On average, an increase of 10 colocation cabinets is related to approximately $9 more profit, 6 more occurrences and mispricing lasts 31 seconds longer in the two markets. The increase in duration of high frequency arbitrage profit opportunities contrasts with the findings of Budish et al [2013] who report a *decline* in the duration of arbitrage profit opportunities over time even as total index arbitrage profits remain essentially constant.

[INSERT TABLE 4]

Kozham and Tham [2012] argue that HFT competition increases execution risk as there may be several HFTs submitting the same trade and adversely affecting pricing efficiency. On the other hand, more HFTs can also result in greater competition for liquidity. With closer monitoring of market conditions, HFTs are able to exploit arbitrage profit opportunities that were unlikely to be realized prior to colocation, due to inconsistency in latency that negatively impacts the probability of execution. Collectively, we observe that the over-competition of HFTs may out-weigh the benefit HFTs bring, which results in a market with larger more frequent and prolonged price discrepancies between the two instruments.

Among the control variables, the lag index return is statistically insignificant for all three model specifications in Panel A. This implies that observed arbitrage opportunities are short-lived and not influenced by general equity market behavior.[19] *SPI* trading volume is positively and significantly related to arbitrage profits and occurrences. At the 1 percent significance level, *Daily profit* and *Duration* are both negatively related with *log(1/SPI Dollar Volume)*: a 1 percent increase is associated with an increase of $3 in arbitrage profit and almost 1 extra arbitrage profit opportunity. Price volatility is positively and significantly related to the size and number of arbitrage opportunities. This is consistent with the extant literature that suggests large trading volume and volatility induce greater future index mispricing, which, in turn, leads to more arbitrage profit opportunities in the market.

[19] For robustness, we also repeated the analysis with the current day index return instead of the previous day value and the result does not differ from what we have reported.

For robustness, we also use an alternative HFT proxy, *log(Message Traffic)* to replace the colocation variable and re-estimate the regression models. HFTs employ low latency market access to closely observe the market and, at the same time, utilize larger amounts of orders to capture any available opportunities. *Message Traffic* captures such information by recording the daily total number of order book messages from both *SPI* and *STW*.[20] Since the variable *Message Traffic* is a continuous measure and is observed over the entire sample period, we are able to utilize the entire sample period (01/03/12-31/01/2014) in our estimates. The results are reported in Panel B of Table 4.

The results based on *Message Traffic* reported are consistent with our earlier findings: both *Daily Profit* and *Duration* yield significantly positive coefficients, implying that when the two markets are inundated with message traffic, more price discrepancies are observed, which leads to more arbitrage opportunities (in terms of both value and frequency).[21] However, the *Message Traffic* variable does not the duration of arbitrage opportunities.

Robustness Test: Incorporating Transaction Cost

Other than the cost of crossing the bid-ask spread, our analysis so far has not considered exchange fess or the cost of co-location. Such costs can be categorized as fixed or variable costs. Fixed costs include the cost of setting up the server (e.g. computer hardware & server space renting), and market data feeds which are not conditioned on participants' trading activities[22]. For any existing brokers, who have already obtained co-location service and operated in the market, fixed costs are a sunk cost and therefore irrelevant. Variable costs, on the other hand, are more crucial to the profitability of the strategy examined in this paper, since they are directly related to participants' level of trading activity, including arbitrage activity. ASX charges two types of variable costs: trade and clearing fees. For ETF instruments, trading and clearing fees are 0.15 (subject to total monthly maximum of $75) and 0.90 basis points for each trade value. Futures trading incur a $1.00 trading fee and $0.90 clearing fee per contract side traded.

We incorporate the variable trading fee structure in identifying arbitrage opportunities.[23] Descriptive statistics of updated arbitrage variables (after transaction cost) are reported in Table 6.

[INSERT TABLE 5]

Not surprisingly, the daily frequency of arbitrage opportunities, profit and duration are substantially less than analogues figures documented earlier, since a large proportion of

[20] The presence of both *Message Traffic* and *1/SPI Dollar Volume* forms the order to trade ratio, which is widely used as a proxy for algorithmic trading. (See Hendershott et al., 2011).
[21] Daily profit and duration, which are larger than 2000 dollars/seconds, are considered as outliers and are excluded from all of the regressions.
[22] It costs $2,500/month for a participant to co-locate in the ALC. To enable the discussed strategy, the participant needs to equip with the lowest latency connections: ASX 24 ITCH for accessing order book information ($6,000/month) and ASX OUCH for executing trades ($6,250/month).
[23] The strategies introduced here are carried out by taking long/short positions upfront and reversing them after a sh few moments. This approach requires round trip trades and hence doubles the variable cost.

arbitrage generates very small profits and are not able to cover the cost of trading. For example, the number of arbitrage opportunities decreases to 24.7, corresponding to a potential daily profit of almost $157

[INSERT TABLE 6]

Table 6 reports regression results based on transaction cost adjusted variables. The results are consistent with our earlier findings. That is, the number of occupied colocation cabinets remain positively correlated with the three arbitrage variables, suggesting that even after transaction costs, an increase in HFT activities is associated with increasing arbitrage opportunities, profitability and duration.

5. Summary and Conclusions

Does competition among high frequency traders reduce the amount of profits available from exploiting fleeting potential arbitrage opportunities? Limits to arbitrage suggest that all apparent profits are unlikely to be entirely eliminated. This study attempts to answer that question, using *duration*, *frequency* and *amount* of arbitrage profit opportunities between the ASX Share Price Index futures contract (*SPI*) and its exchange traded fund (ETF) counterpart, *STW*.

The sharp growth in expensive collocated cabinets on the ASX suggests that HFTs expect to earn substantial profits from exploiting various low-latency trading strategies. Our results suggest that index arbitrage is not likely to be a profitable strategy.

We find that, as the competition among HFTs increases, the negative impacts they bring along outweigh the benefits. Rising occupied colocation cabinets is significantly related with an increase in mispricing and index arbitrage opportunities in all three aspects: value, frequency and duration. With their low latency trading strategies, HFTs collectively have the potential to significantly alter bid and ask prices in a short period of time. This, in turn, creates greater misalignment in prices between two closely related instruments. However, in sharp contrast to what Budish et al [2013] report for the U.S. market, the total amount of gross profit available to an HFT dedicated solely to exploiting index arbitrage profit opportunities between stock index futures and the corresponding ETF in the Australian market is trivial. It would barely cover the costs imposed by the ASX for colocation services let alone the technology costs associated with low latency trading. Moreover, this assumes that the HFT would capture **all** of the potentially available arbitrage profit opportunities. Consistent with what Miller, Muthuswamy and Whaley [1994] report, one interpretation of this result is that there is simply less high frequency index arbitrage in the Australian market than what the findings of Budish et al for the American market would lead one to expect exists in other markets. One possible explanation is that the less actively traded ETF is simply priced off of the corresponding futures eliminating many potential arbitrage profit opportunities in the Australian market. This result of limited index arbitrage profits is also consistent with the industry estimate of declining profits for equity HFTs in the U.S.A. which Tabb [2014] argues has resulted in lower trading costs to the benefit of institutions and other investors.[24]

The fact that index arbitrage is not profitable to a new HFT entrant in the Australian market does not explain why existing HFTs who engaged in other trading strategies do not attempt to capture the (albeit small) index arbitrage profit opportunities. While the study shows that the conventional theory of competition improving market efficiency regardless of time scale does not seem to be effective among HFTs in Australia that does not necessarily mean that competition has no impact on HFTs and market arbitrage profit opportunities. Rather, our evidence may suggest that the presence of HFTs in the Australian market is not at an efficient level. If there had been a sufficient number of HFTs in the market, competitions among them would have eliminated most, if not all, potential arbitrage profit opportunities. The fact that we observe more (albeit small) arbitrage opportunities (even after including transaction cost) with a greater presence of HFT's, indicates that there is the potential for more HFTs to participate in the market and eliminate such opportunities. This applies especially to existing HFTs who are currently engaged in Either some of the current participants modify their strategies and improve their ability to capture these currently unexploited profit opportunities, or new HFTs should establish operations in Australia market to do so. Until then, adverse competition among HFTs may continue to create more market distortions and price discrepancies.

[24] For example, Larry Tabb [2014] argues: "Looking at recent public data, the profitability of HFT firms in the US equities market has declined, just as the number of players has decreased. If the exchanges, brokers and HFTs are not reaping the rewards, then where is this leakage going? This money is going back to investors in the form of better and cheaper executions, as few if any institutional investors we have interviewed – and we have interviewed thousands – have ever expressed that their equity implementation costs have increased, meaning … trading just becomes cheaper and cheaper. That cost comes from somewhere: market makers, speculators, brokers and exchanges. …"

References

Baron, M., Brogaard, J., and Kirilenko, A. (2012). The trading profits of high frequency traders”. Tech. rep., Commodity Futures Trading Commission. Working paper.

Brailsford, T. and Hodgson, A. (1997). Mispricing in stock index futures: A re-examination using the SPI, Australian Journal of Management, 22 (1), 21-45.

Budish, E., Cramton, P., and Shim, J. (2013). The high-frequency trading arms race: Frequent batch auctions as a market design response. Manuscript, 6.

Chaboud, A., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2013). Rise of the machines: Algorithmic trading in the foreign exchange market. Journal of Finance, Forthcoming.

Cummings, J. R., & Frino, A. (2011). Index arbitrage and the pricing relationship between Australian stock index futures and their underlying shares. Accounting & Finance, 51(3), 661-683.

Frino, A., and McKenzie, M. (2002). The pricing stock index futures spreads at contract expiration. Journal of Futures Markets, 22, 451-469.

Frino, A., Mollica, V. and Webb, R. I. (2014). The Impact of Co-Location of Securities Exchanges' and Traders' Computer Servers on Market Liquidity. Journal of Futures Markets, 34, 20–33.

Hendershott, T., Charles M. J., and Albert J. M. (2011). Does Algorithmic Trading Improve Liquidity?. The Journal of Finance 66, 1-33.

Kozhan, R., and Tham, W. W. (2012). Execution risk in high-frequency arbitrage. Management Science, 58, 2131-2149.

Newey, W. K., West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55, 703–708.

Scholtus, M., van Dijk, D., and Frijns, B. (2014). Speed, algorithmic trading, and market quality around macroeconomic news announcements. Journal of Banking & Finance, 38, 89-105.

Miller, M. H., Muthuswamy, J., and Whaley, R. E. (1994). Mean reversion of Standard & Poor's 500 Index basis changes: Arbitrage - induced or statistical illusion? The Journal of Finance, 49(2), 479-513.

Larry Tabb, “No, Michael Lewis, the US Equities Market Is Not Rigged,” TABB Group, March 31, 2014, available at http://tabbforum.com/opinions/no-michael-lewis-the-us-equities-market-is-not-rigged?print_preview=true&single= true.

Webb, R. I. (2003). Transitory real-time property rights and exchange intellectual property. Journal of Futures Markets, 23, 891–913.

Table 1: Descriptive Statistics of *SPI* and *STW*

This table reports descriptive statistics for the period March 1,2012 to January 31, 2014 (Panel A) and five event windows (Panel B). *Spreads*, measure in index points for *SPI*, cents for *STW*, and best level *Depth* (both converted in dollar) are time weighted. Number of trades (*# Trades*), number of message updates (*message traffic*) and *Dollar Volume* are measured as at the end of each trading day. *Volatility* is evaluated by taking the log of highest and lowest mid-quote price of the trading day, expressed in percentage.

	Spread (Ind pts/cents)		*Depth ($1000)*		*# Trades*		*Dollar Volume ($Million)*		*Message Traffic (000's)*		*Volatility (%)*	
	SPI	**ETF**	**SPI**	**ETF**	**SPI**	**ETF**	**SPI**	**ETF**	**SPI**	**ETF**	**SPI**	**ETF**
Panel A: Summary Statistics												
March 2012 to March 2014, 480 trading days included												
Mean	1.130	2.243	4675	1758	8243	335	1791	8.344	74.50	12.84	0.764	0.762
Std. Dev.	0.060	0.431	1227	723	2192	214	543	6.858	18.23	5.96	0.339	0.354
Q1	1.085	1.959	3830	1269	6738	196	1430	4.621	60.76	8.79	0.515	0.507
Median	1.119	2.215	4464	1662	7893	288	1697	6.171	73.58	12.00	0.692	0.686
Q3	1.164	2.461	5317	2196	9175	406	2059	9.625	84.85	15.91	0.930	0.938
Panel B: Event Window Summary												
Window 1(March 2012, 20 co-located cabinets, 22 trading days included)												
Mean	1.130	2.345	5179	1882	9206	239	1906	4.840	83.17	8.93	0.701	0.691
Std. Dev.	0.033	0.342	732	485	1792	87	414	3.255	10.13	4.40	0.222	0.220
Window 2(May – July 2012, 76 co-located cabinets, 65 trading days included)												
Mean	1.089	2.237	4614	2504	9178	276	1769	11.658	78.27	19.66	0.892	0.905
Std. Dev.	0.030	0.603	626	856	2371	142	522	11.221	9.38	3.83	0.377	0.411
Window 3(November 2012 – January 2013, 111 co-located cabinets, 60 trading days included)												
Mean	1.082	1.992	6327	1585	7539	357	1611	8.855	72.34	19.26	0.603	0.590
Std. Dev.	0.028	0.312	1110	502	1849	199	438	6.852	10.62	3.17	0.239	0.236
Window 4(May – July 2012, 117 co-located cabinets, 65 trading days included)												
Mean	1.208	2.375	3534	1329	8909	363	2049	9.226	76.83	19.78	0.965	0.977
Std. Dev.	0.052	0.469	591	383	2592	161	679	6.266	19.45	5.72	0.429	0.442
Window 5(November 2013 – January 2014, 133 co-located cabinets, 59 trading days included)												
Mean	1.151	2.310	3868	1967	8266	335	1840	8.121	66.21	16.98	0.806	0.81
Std. Dev.	0.052	0.394	826	811	2102	177	530	4.311	17.05	6.11	0.322	0.322

Table 2 Arbitrage Summary Statistics (March 1, 2012 – January 31, 2014)

This table reports the mean and percentiles of arbitrage variables from executing the trading strategy between the *SPI* and *STW*, as described in section 2. *# of Arbs/Day* records counts of arbitrages for each trading day. *Qty*. denotes the size of each arbitrage opportunity, measured in equivalence to the number of SPI contract traded. Per-Arb Profits measures the actual dollar amount of each arbitrage, which multiplies the difference between *SPI* price and 100 times *STW* price, at the start of each opportunity, with the minimum volume available to both instruments. Accumulating *Per-Arb Profits ($)* at the end of each trading day yields *Daily Profits*. *Daily Duration* accumulate the time that each individual arbitrage opportunity lasts over each trading day.

Convergence Time records the time used in each trading day to calculate the basis spreads between *SPI* and *STW*, and described as r in section 2. *% SPI initiated* records the percentage of arbitrage opportunities that are initiated by a change in the price of SPI. *% Good Arbs* presents the percentage of arbitrage reverted before the daily convergence time. % Buy vs. Sell reveals the proportion of arbitrages that require the bid strategy (long *SPI* & short *STW*).

		Percentile						
	Mean	*1*	*5*	*25*	*50*	*75*	*95*	*99*
# of Arbs/Day	124	34	54	81	114	154	236	355
Qty (SPI Lots)	1.126	0.001	0.004	0.040	0.258	1.000	5.000	9.950
Daily Profit $	529	52	101	244	430	671	1342	2312
Daily Duration	606	37	94	255	451	762	1582	3110
Convergence Time (sec)	75.36	16.02	20.67	40.90	65.04	106.94	150.97	164.85
% SPI Initiated	81.88%							
% Good Arbs	98.15%							
% Buy vs. Sell	50.64%							

Table 3 Regression Variables Summary Statistics (for all event windows)
This table shows the mean and standard deviation (in parenthesis) of the arbitrage and control variables, used in the regression analysis. Panel A contains the arbitrage variables from executing trading strategy between the *SPI* and *STW*, as described in section 2. *Daily Profits* and *Daily Duration* accumulate the profit and lasting time of individual arbitrage opportunities at the end of each trading day respectively. *No. of Arbitrage* shows the total number of arbitrage opportunities during each window.
Panel B contains control variables which has not been presented in the previous tables. *Colo. Cabinets* records the number of cabinets reported at each time period. *Log(Message Traffic)* is the log of daily order book message counts. *Index Return* records the daily index return. *News Release* captures total macro-news relevance level, recorded by Bloomberg.

	All ***01/03/12-31/01/14***	***Window 1*** ***01/03/12-31/03/12***	***Window 2*** ***01/05/12-31/07/12***	***Window 3*** ***01/11/12-31/01/13***	***Window 4*** ***01/05/13-31/07/13***	***Window 5*** ***01/11/13-31/01/14***
Panel A: Arbitrage Variables						
Daily Duration	606.1	232.9	489.6	681.5	370.4	749.9
(sec)	(551.9)	(117.3)	(375.7)	(372.5)	(278.5)	(543.4)
Daily Profit	528.6	319.9	681.1	352.9	611.9	639.7
($)	(416.3)	(145.4)	(476.9)	(348.3)	(386.9)	(544.8)
No. of	124.4	77.2	104.0	100.1	156.5	147.3
Arbitrage	(60.4)	(25.6)	(42.2)	(34.8)	(77)	(56.8)
Panel B: Control Variables						
Colo. Cabinets		20	76	111	117	133
Log(Message	11.34	11.44	11.35	11.29	11.44	11.30
Traffic)	(0.232)	(0.129)	(0.245)	(0.261)	(0.253)	(0.266)
Index Return	-0.019	0.038	-0.050	0.022	-0.173	-0.095
(%)	(0.506)	(0.404)	(0.62)	(0.377)	(0.598)	(0.543)
News	100.2	106.7	106.5	107.8	104.2	98.9
Release	(124.7)	(147.4)	(131.5)	(128.2)	(127.7)	(124.9)

Table 4: Regression Analysis of Arbitrages

Panel A of the table presents regression analysis of the impact of HFT completion on market arbitrage, with data in the five event windows. The following specification is estimated:

$$Arb_t = Intercept + \beta_1 Colocation_t + \beta_2 \, log(1/ \quad SPI\ Dollar\ Volume_t) + \beta_3 Volatility_{SPI,\,t} + \beta_4 NewsRelevence_t + \beta_5 LagReturn_t,$$

where Arb_t is any of the three dependent variables: *Daily Profit*, *Daily Duration* or *# Arbitrage*. $Colocation_i$ is the number of cabinets hosted during the event window and our proxy for HFT competition. $log(1/SPI\ Dollar\ Volume)$ is log of the inverse dollar value of traded *SPI* contracts on day i; $Volatility_{SPI,i}$ measures the daily volatility level in *SPI*, using log of the ratio between the highest and lowest mid-price of SPI contract; $NewsRelevence_i$ captures total macro-news relevance level, recorded by Bloomberg, on day *i*; $LagReturn$ is the return of S&P/ASX200 index in the previous trading day.

Panel B adjust the regression by substituting the key variable $Colocation_i$ with $log(Message\ Traffic_i)$, which records the log of daily number of order book message in both *SPI* and *STW*. All other control variables remain the same, and data of the entire sample period (01/03/12~31/01/14) are used. *t*-statistics are reported in parenthesis and Newey-West (1987) adjusted.

	Panel A			***Panel B***		
	Daily Profit	*No of Arb*	*Daily Duration*	*Daily Profit*	*No of Arb*	*Daily Duration*
Intercept	-6704***	-1724***	-949	-5430***	-1605***	-2989
	(-3.81)	(-6.44)	(-0.43)	(-4.09)	(-6.26)	(-1.44)
Colocation	0.928*	0.652***	3.08***			
	(1.7)	(8.87)	(6.39)			
Log(Message Traffic)				193**	36.3***	-160
				(2.55)	(2.65)	(-1.49)
Log(1 / SPI dollar volume)	-324***	-82.8***	-57.1	-165*	-60.8***	-254**
	(-3.84)	(-6.45)	(-0.54)	(-2.15)	(-4.30)	(-2.12)
Volatility	30805***	3188***	-3978	32430***	3713***	-5967
	(4.34)	(2.74)	(-0.49)	(5.89)	(3.59)	(-0.78)
New Release	-0.137	-0.045**	-0.298**	-0.119	-0.039**	-0.172
	(-0.96)	(-2.34)	(-2.04)	(-1.09)	(-2.24)	(-1.20)
Lag Return	918	195	-1801	2221	8.895	-4851
	(0.22)	(0.32)	(-0.46)	(0.74)	(0.02)	(-1.28)
Adjusted R^2	0.259	0.424	0.074	0.240	0.293	0.006

*** significant at 1%
** significant at 5%
* significant at 10%

Table 5: Arbitrage Summary Statistics after trading and clearing expense
(March 1, 2012 – January 31, 2014)

This table shows the mean and percentiles of arbitrage variables from executing trading strategy between the *SPI* and *STW*, as described in section 2, after adjusted for cost of trading. *# of Arbs/Day* records daily counts of arbitrages for each trading day. *Qty.* denotes the size of each arbitrage opportunity, measured in equivalence to the number of SPI contract traded. Per-Arb Profits measures the actual dollar amount of each arbitrage, which multiplies the difference between *SPI* price and 100 times *STW* price, at the start of each opportunity, with the minimum volume available to both instruments. Accumulating Per-Arb Profits at the end of each trading day yields *Daily Profits*. *Daily Duration* accumulates the lasting time of each individual arbitrage opportunity at the end of each trading day.

		Percentile						
	Mean	*1*	*5*	*25*	*50*	*75*	*95*	*99*
# of Arbs/Day	24.7	3	6	14	21	33	57	79
Qty (SPI Lots)	1.13	0.00	0.00	0.04	0.26	1.00	5.00	9.95
Daily Profit ($)	156.53	3.02	10.27	45.21	101.11	197.72	497.71	851.73
Daily Duration (sec)	4.678	0.0042	0.0146	0.0742	0.311	2.216	27.252	61.766

Table 6: Regression Analysis of Arbitrages (after trading and clearing expense)
Arbitrage variables (*Daily Profit*, *Daily Duration* or *# Arbitrage*) are updated with expense adjusted arbitrage profit. Panel A of the table presents regression analysis of the impact of HFT completion on market arbitrage, with data in the five event windows. The following specification is estimated:

$$Arb_i = Intercept + \beta_1 Colocation_i + \beta_2 \log(1/SPI\ Dollar\ Volume) + \beta_3 Volatility_{SPI,i} + \beta_4 NewsRelevence_i + \beta_5 LagReturn_i,$$

where Arb_i is any of the three dependent variables: *Daily Profit*, *Daily Duration* or *# Arbitrage*. $Colocation_i$ is the number of cabinets hosted during the event window and our proxy for HFT competition. $log(1/SPI\ Dollar\ Volume)$ is log of the inverse dollar value of traded *SPI* contracts on day i; $Volatility_{SPI,i}$ measures the daily volatility level in *SPI*, using log of the ratio between the highest and lowest mid-price of SPI contract; $NewsRelevence_i$ captures total macro-news relevance level, recorded by Bloomberg, on day i; $LagReturn$ is the return of S&P/ASX200 index in the previous trading day.
Panel B adjust the regression by substituting the key variable $Colocation_i$ with $log(Message\ Traffic_i)$, which records the log of daily number of order book message in both *SPI* and *STW*. All other control variables remain the same, and data of the entire sample period (01/03/12~31/01/14) are used. *t*-statistics are reported in parenthesis and Newey-West (1987) adjusted.

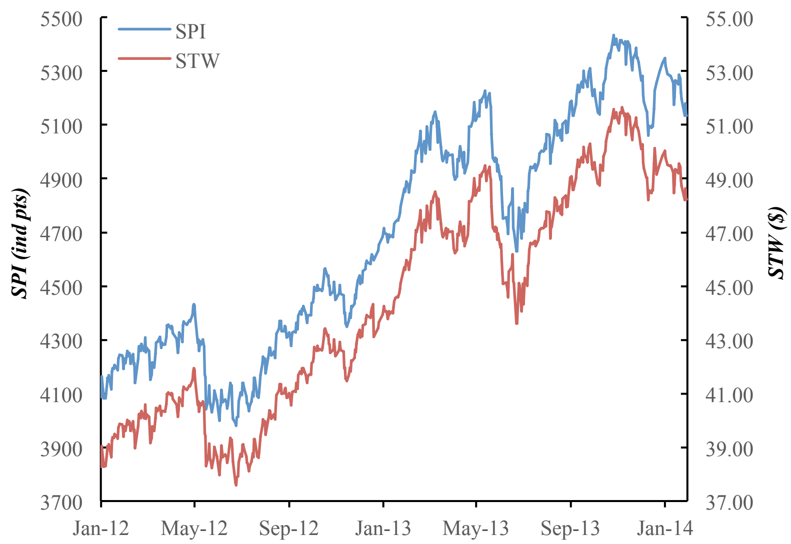
	Panel A			***Panel B***		
	Daily Profit	*No of Arb*	*Daily Duration*	*Daily Profit*	*No of Arb*	*Daily Duration*
Intercept	-955.0	-407.8***	-232.1	-888.3	-326.3***	-222.3
	(-1.09)	(-5.69)	(-0.39)	(-1.00)	(-4.47)	(-0.39)
Colocation	0.451*	0.089***	0.522***			
	(1.83)	(4.41)	(3.08)			
Log(Message Traffic)				-0.672	17.381***	-18.083
				(-0.01)	(3.68)	(-0.43)
Log(1 / SPI dollar volume)	-46.27	-19.78***	-12.30	-45.58	-7.04	-23.95
	(-1.11)	(-5.73)	(-0.43)	(-0.83)	(-1.50)	(-0.67)
Volatility	11226***	443	2863	11385***	548	2973
	(3.03)	(1.16)	(0.98)	(3.08)	(1.44)	(1.04)
New Release	-0.124*	-0.018***	-0.101**	-0.125*	-0.018***	-0.103**
	(-1.85)	(-3.42)	(-2.39)	(-1.85)	(-3.35)	(-2.37)
Lag Return	53.68	-14.88	350.31	-124.4	-51.5	143.2
	(0.03)	(-0.08)	(0.20)	(-0.06)	(-0.27)	(0.08)
Adjusted R^2	0.101	0.277	0.037	0.093	0.275	0.014

*** significant at 1%
** significant at 5%
* significant at 10%

Figure 1: Price Path of SPI and STW

This figure displays the price path over the entire sample period (March 2012 – January 2014), Panel A, and a sample one-hour snapshot on a randomly selected trading day 28/06/2013, Panel B. Prices plotted in Panel A are the mid quote price at 4:00pm from *SPI* and *STW* on each trading day. Prices in Panel B are the mid quote price, each time the best bid and offer prices changed in wither *SPI* and *STW*, during the selected window.

Panel A: March 2012 – January 2014



Panel B: 2:00 PM – 3:00 PM, June 28th, 2013

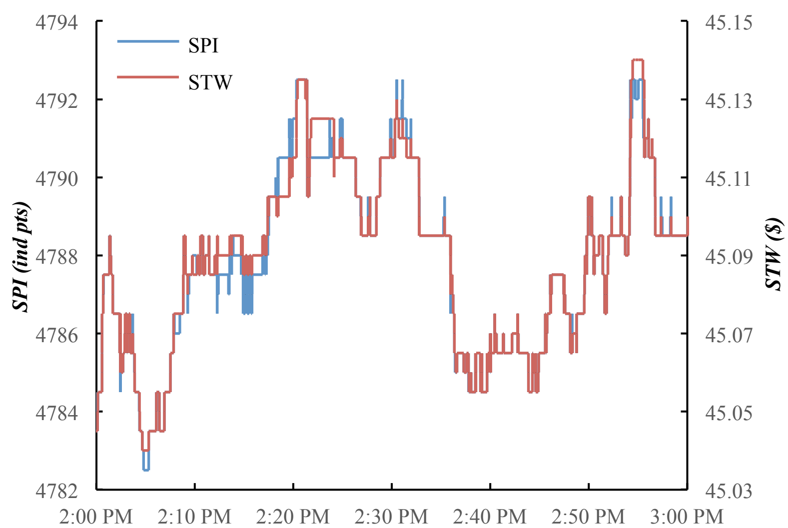
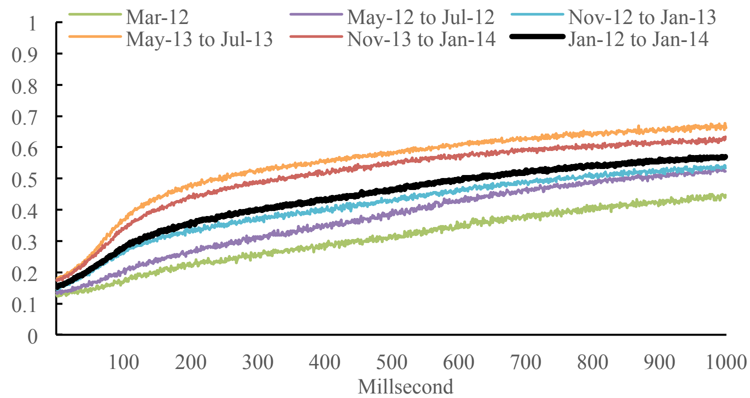


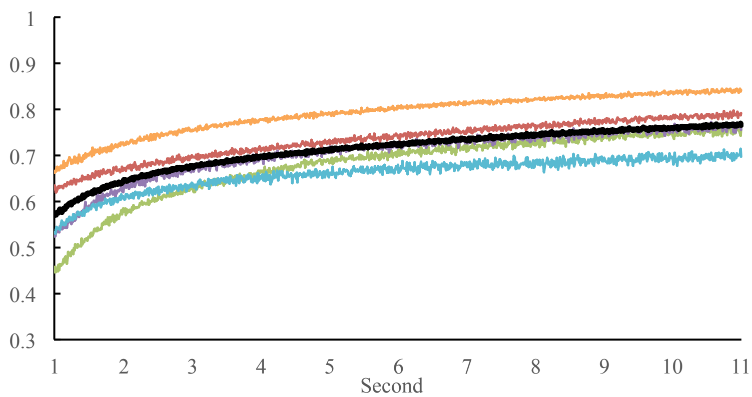
Figure 2: Median Correlation convergence of the sample and events

This figure displays median daily correlation using mid-price return of the S&P/ASX 200 index future, *SPI*, and the index ETF, *STW*, sampled at various time gaps, over the entire sample period of Jan 2012 – Jan 2014 and five event windows. Sampled price is calculated using time weighted average in each time gap. Each point is the median correlation over all trading days during the period or event windows. Panel A presents returns sampling from 1 millisecond gap to 1000 milliseconds and Panel B shows the correlation of time gaps from 1 second to 11 second with increases per 10 milliseconds.

Panel A: 1 to 1000 milliseconds, per millisecond increment



Panel B: 1 to 11 seconds, per 10 millisecond increment



Panel C: 11 to 111 seconds, per 100 millisecond increment