HONG KONG INSTITUTE FOR MONETARY RESEARCH

ORDER IMBALANCE AND THE PRICING OF INDEX FUTURES

Joseph K.W. Fung

HKIMR Working Paper No.13/2006

October 2006

Hong Kong Institute for Monetary Research

(a company incorporated with limited liability)

All rights reserved.
Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

Order Imbalance and the Pricing of Index Futures

Joseph K.W. Fung
Hong Kong Baptist University
Hong Kong Institute for Monetary Research

October 2006

Abstract

This study examines whether the direction and magnitude of the aggregate order-imbalance of the index stocks can explain the arbitrage spread between index futures and the underlying cash index. The data are for the Asian financial crisis period and hence entail wide variations in order imbalance and the index-futures basis. The analysis controls for realistic trading costs and actual dividend payments. The results indicate that the arbitrage spread is positively related to the aggregate order imbalance in the underlying index stocks; negative order-imbalance has a stronger impact than positive order imbalance. Violations of the upper no-arbitrage bound are related to positive order imbalance and violations of the lower no-arbitrage bound are related to negative order imbalance. Asymmetric response times to negative and positive spreads can be attributed to the difficulty, cost, and risk of short stock arbitrage when the futures is below its no-arbitrage value. The significant relationship between order imbalance and arbitrage spread confirm that index arbitrageurs are important providers of liquidity in the futures market when the stock market is in disequilibrium.

Acknowledgements: Part of this study was undertaken when Joseph Fung was a Research Fellow at the Hong Kong Institute for Monetary Research. He acknowledges the financial support of a faculty research grant from Hong Kong Baptist University (FRG/00-01/II-28) and of an earmark research grant provided by the Research Grants Council of Hong Kong (HKBU 2013/02H). The author thanks Kin Lam for many helpful comments and suggestions; Hans Genberg, Matthew Yiu, seminar participants at Hong Kong Baptist University, and the Hong Kong Institute for Monetary Research, and Thomas Kong and Agnes Lee for research assistance.

The views expressed in this paper are those of the author, and do not necessarily reflect those of the Hong Kong Institute for Monetary Research, its Council of Advisors, or the Board of Directors.

1. Introduction

Arbitrage opportunities contradicts the fundamental notion of an efficient capital market. It is generally argued that apparent index arbitrage opportunities due to misalignments between index futures and the underlying index reflect market frictions (Kawaller, 1987), or are simply statistical illusions (Miller, Muthuswamy, and Whaley, 1994). Yet authors continue to provide empirical evidence of index arbitrage opportunities, and a significant portion of stock trades are associated with index arbitrage activities.[1]

Harris, Sofianos, and Shapiro (1994) report that about 50% of the 50,760 NYSE program trades between January 1989 and December 1990 were related to index arbitrage. Index arbitrage is a substantial business, and its potential profit must be a rent to compensate arbitrageurs for their unique services.

Grossman writes: "If there is a large net supply of futures because institutions want to sell stocks, then this will drive the futures price down....Index arbitrageurs will view this differential as a profit opportunity. They will buy futures and sell common stock. They are the natural buyers of futures in situations where futures are being sold because investors, in net, want to hold less common stock" (Grossman, 1988, p. 291).

Following Grossman's conjecture, when there is excess selling of stocks, stock futures should be priced at a discount relative to the "fair" or no-arbitrage value, in order to create a buy futures-short stock opportunity, to motivate the supply of short futures from arbitrageurs. Similarly, the futures should be priced at a premium over its fair value, in order to create a short futures-buy stock arbitrage opportunity, to motivate the supply of long futures by arbitrageurs when there is excess buying in the stock market.

In times of excess selling (buying) pressure in the stock market, both hedger and speculator demands for short (long) futures positions increase. In these cases, the observed premium or discount should represent real potential arbitrage opportunities to index arbitrageurs. Hence, the observed premium or discount should be potentially exploitable by index arbitrageurs to compensate them for supplying liquidity in the futures market.

My research provides an empirical test of this proposition. Following Blume, MacKinlay, and Terker (1989) (BMT), I use measures of order imbalance as proxies for the state of excess selling or buying in the cash index stocks. Order imbalance is defined as buyer-initiated trading volume (trades executed at ask prices) minus seller-initiated trading volume (trades executed at bid prices). When the market is in equilibrium, transactions should be conducted randomly at either the bid or the offer price, and the volumes of stock traded at the two prices would be roughly equal. When there is a large positive order imbalance in the index stocks, this indicates excess buying in the cash market; a large negative order imbalance indicates excess selling.

1 See for examples Sofianos (1993) and Neal (1996) for the U.S. markets; Yadav and Pope (1990, 1994) for the U.K. markets; and Fung and Draper (1999) and Draper and Fung (2003) for the Hong Kong markets.

Roll, Schwartz, and Subrahmanyam (2005) find that the absolute level of the index-futures basis is positively related to the illiquidity of the market. They use both quote and effective spreads as proxies for market liquidity. Order imbalance, also a proxy for the liquidity of the market, is a measure of the extent of liquidity asymmetry.

Positive order imbalance indicates liquidity for the sell side but relative illiquidity for the buy side, because buying is more active than selling. A demand for liquidity from the buy side could spill over to the related futures market; increases in buying pressure on index futures cause the positive basis to widen. Similarly, negative order imbalance indicates poor liquidity for the sell side but improved liquidity for the buy side. The demand from the sell side for liquidity induces greater demand for short futures positions, and drives the basis negative.

Order imbalance may also be a proxy for a signaling effect on the futures price relative to the cash index. Positive order imbalance may signal a rise in the cash market, when traders may buy futures ahead of an impending stock price movement. Similarly, negative order imbalance may signal a potential drop in the market, and traders will short futures ahead of the cash market decline. The signaling effect should also induce a positive relation between order imbalance and the futures-index basis. Hence, the signaling effect reinforces the liquidity effect.

Studies of index arbitrage opportunities are potentially plagued by infrequent trading and a bid/ask bounces in the index constituent stocks. Thus there are significant measurement problems in identifying prospective transaction (buying and selling) prices of the index basket and in measuring the potential profitability of arbitrage.

Following Draper and Fung (2003), I attempt to address the problem by using bid and ask index prices that are estimated from synchronous and active stock quotes retrieved from a screen-based trading system. The quotes in the system are firm commitments, meaning the measured arbitrage opportunities are forward-looking and potentially executable. Use of quotes and transactions data from a screen-based trading system also allows for accurate identification of seller-versus-buyer initiated trades since the system eliminates the possibility of trade-throughs.

To mitigate the effect of dividend payments on the index component stocks, I discount the actual (ex-post) dividend payments. I also assume realistic trading costs in identifying arbitrage opportunities.

The results show that arbitrage spreads are positively related to aggregate order imbalance of index stocks. Negative order imbalance has a stronger influence than positive order imbalance on arbitrage spreads, reflecting the difficulties, costs, and risks of short-selling stocks for marginal arbitrageurs. These findings are consistent with the proposition that index arbitrageurs provide liquidity in the futures market, especially when the stock market is in disequilibrium. The results also suggest that it is less costly and easier for the market to resolve liquidity problems for the buy side than for the sell side, again because of the risk, cost, and inconvenience associated with taking short stock positions.

2. Literature review

Research in the U.S. markets has indicated futures are usually underpriced relative to the underlying index. Modest and Sundaresan (1983) attribute the discount to the inability of marginal arbitrageurs to fully use the proceeds from short-selling of the index portfolio. Chan (1992) and Neal (1996) argue that short-selling should not matter because institutional investors can practice quasi-arbitrage.

Kempf's (1998) study of the German DAX contract, and some European markets suggests that the costs of short-selling affect the mispricing of futures contracts. Studies of the Hang Seng Index contracts by Fung and Draper (1999), Fung and Jiang (1999), and Jiang, Fung, and Cheng (2001) show that the lifting of restrictions against short-selling reduces the size and the frequency of discounts, and speeds up the dynamic adjustment process, especially when futures are underpriced. Draper and Fung (2003) find that a large negative basis in Hang Seng Index futures is associated with the lack of short-selling in the index constituent stocks following market intervention by the Hong Kong government.

Cornell and French (1983) conjecture that the presence of a tax timing option in stocks induces a premium in the cash index over the futures. Bhatt and Cakici (1990) show, however, that futures could be priced at a premium.

All these studies suggest that market conditions can affect the direction and extent of mispricings, but Yadav and Pope (1994) do not find a significant relationship between (rising and falling) market conditions and mispricing. Chan, Chung, and Johnson (1993) do show that increases in market volatility reduce the arbitrage spread. Chen, Cuny, and Haugen (1995) report that market conditions affect the customization value of stocks relative to futures that determine the index-futures basis. An increase in market volatility reduces the customization value of stock, which lowers the basis.

If all this research helps to explain why some particular mispricing persists, the driving force behind the mispricing remains unclear. And expanding upon this line of research, and testing whether order imbalance in the stock market can explain the premium or discount in the index futures price.

Other authors examining the relation between order imbalance and the stock market volatility include Locke and Sayers (1993) and Chan and Fong (2000). Locke and Sayers (1993) define order imbalance as the absolute value of the difference between all contracts bought and sold within a minute by floor traders for their own account; they study the relationship of the return variance and order imbalance together with other variables that proxy for information arrival. Chan and Fong (2000) show that the volume-volatility relation could be explained by the order imbalance (defined as the net of the numbers of buyer-versus-seller initiated trades). They find that, on a daily basis, order imbalance is highly correlated with the total number of trades in both NYSE and NASDAQ stocks; the volume-volatility relation is weaker after capturing the impact of order imbalance on the intraday stock return. Still, these authors do not explore whether the direction of price movement is related to the sign of the order imbalance.

Blume, MacKinlay, and Terker (1989) (BMT) define order imbalance as the difference between the dollar volume crossed at asked and crossed at bid. Correlations between the aggregate order imbalance and the concurrent 15-minute market returns were at 0.81 and 0.86 on October 19 and 20. The correlations are significant at the individual stock level.

Rather than measure order imbalance in the stock market, Easley, O'Hara, and Srinivas (1998) differentiate between executed buy and sell orders on call and put options traded on the Chicago Board Options Exchange. They examine whether they together convey information on price movements of the underlying stock. Their results show that executed order flows on a financial product help explain the price movement in related securities.

I adopt the BMT measure of order imbalance to examine how and whether order imbalance in the cash stock market affects the arbitrage spread between the futures and the underlying index. The results shed light on whether order imbalance in the cash stock market is responsible for misalignment between the futures and the cash index.

Order imbalance affects the index and index prices, as well as the mispricing in two possible ways. The first is through the liquidity effect; institutions (or hedgers) compensate arbitrageurs for providing liquidity in the futures market by paying a premium (or accepting a discount) when positive (negative) order imbalance in the cash stock market forces the institutions to substitute cash positions for futures.

Second is the signaling effect; speculators buy and sell futures contracts, taking positive order imbalance as a signal of an impending rise in the stock market and negative order imbalance as a signal of a fall. In this respect, the liquidity and the signaling effects reinforce each other.

This conjecture is consistent with findings that changes in futures price usually lead changes in the underlying index (e.g., Kawaller, Koch, and Koch (1987) and Stoll and Whaley (1990) for the U.S. market; and Fung and Jiang (1999), and Jiang, Fung, and Cheng (2001) for the Hong Kong market).

Studies of the index-futures arbitrage relation have been plagued by measurement problems caused by infrequent trading and the bid/ask bounce. To alleviate these effects, I use a reconstructed time series of the index based on highly synchronous and active bid/offer quotes of the constituent stocks.[2] The data also reduces the potential distortion of execution and reporting lags that affects studies using data obtained from floor trading systems.

The underlying index for the futures contract includes only 33 stocks, and dividend payments are clustered in the first and the last quarters of the year. The study controls for the impact of dividend payments on the index-futures relation by discounting the cash index for actual (ex-post) dividends accruing to the index.

Chan, Chung, and Johnson (1993) also show that using mid-quotes mitigates the impact of discreteness of tick size on responsiveness of the traded price. Harris (1989), Kleidon (1992), and Kleidon and Whaley (1992) have examined the large negative basis between the S&P 500 index and futures during the U.S. market crash on October 19, 1987.

---

2 See Draper and Fung (2003) for details on the construction of the bid and ask index prices.

Harris shows that the high basis cannot be entirely explained by non-synchronous trading in the stock market. Kleidon and Whaley (1992) argue that the delinkage between the stock and futures markets could have been caused by the inefficient order routing system in the NYSE at that time. Blume, MacKinlay, and Terker (1989), however, find a significant positive relation between order imbalance and the concurrent 15-minute market returns on October 19 and 20. These results seem to suggest that the behavior of the S&P 500 index-futures basis on October 19 could be associated with the pattern of order imbalance. Shleifer and Vishny (1997) show theoretically that widening of the arbitrage basis under extreme market conditions could paralyze arbitrage because of exhaustion of arbitrage capital.

Draper and Fung (2003) examine the behavior of the arbitrage basis during the Hong Kong financial crisis, and find that the index and futures prices remained closely aligned until the Hong Kong government intervened in both the stock and the index derivatives markets. Yet Harris, Sofianos, and Shapiro (1994) find that program trading activities are positively related to market volatility. Hence, it is expected that arbitrage-related trading will intensify during the crisis period, and traders should respond more quickly to mispricing signals.

Breen, Hodrick, and Korajczyk (2002) find that firm-specific characteristics affect the (positive) relation between order imbalance and stock returns. Roll, Schwartz, and Subrahmanyam (2005) find a significant bi-directional causality relationship between market liquidity and the NYSE composite index-future basis. They use both quotes and effective spreads as liquidity measures, and focus on the daily absolute level of the index-futures basis. They find a strong contemporaneous relation between liquidity and the absolute basis. Moreover, the time for the basis to revert to zero is positively related to market liquidity.

Expanding on this research, and examine using both interday and intraday data the relation between arbitrage basis and the aggregate order imbalance of all index stocks. The hypotheses are as follows:

H1: Positive pricing error (i.e., the futures are overpriced) is associated with positive order imbalance, and negative pricing error (i.e., the futures are underpriced) is associated with negative order imbalance. The (signed) pricing error is positively related to order imbalance.

H2: Negative order imbalance has a greater impact than positive order imbalance on the pricing error.

H3: Mispricing disappears more slowly for negative errors than for positive errors.

H4: The time it takes for mispricing to disappear is positively related to the extent of the order imbalance.

H5: The time it takes for a positive error to disappear is positively related to order imbalance, and the time it takes for a negative error to disappear is negatively related to order imbalance.

3. Data and methodology

Time-stamped bid/offer quotes for the 33 constituent stocks of the Hang Seng Index (HSI) and transaction records for stocks and the Hang Seng Index futures are obtained from the Hong Kong Exchange and clearing for May 1, 1996, through April 30, 1998. Dividend information including the ex-date, the payment day, and the actual amount of dividend for the constituent stocks is also obtained from the Exchange.

Stocks were traded electronically via a screen-based Automatic Matching System (AMS) system, and the futures were traded in the pit via the open-outcry method.[3]

The sample period includes the extreme market conditions during the Asian financial crisis. Hence, the data show large intraday and interday variations in stock and futures prices and trading volumes.

To construct the market value weight for each index stock, we obtain market capitalization information and closing index quotes from Hang Seng Index Services Limited. Hong Kong Inter-Bank Offer Rates (HIBORs) for maturities of one-day to one-month are retrieved from Datastream and are taken as the riskless rate of interest.

I generally follows Lee and Ready's (1991) approach to identify seller-initiated (i.e., bid) versus buyer-initiated (i.e., ask) trades. First, a trade is identified according to the nearest previous bid/ask quotes. A trade is identified as a bid (an ask) trade if the traded price is below (above) the mid-point of the quotes. If this measure does not allow identification of a trade, we use the nearest quotes following the trade.

The reason is that the HKEx retrieves quotes by taking snapshots of the limit order book every 30 seconds. Hence, the quote following the trade could have been the quote at which a trade was executed.

When the traded price falls exactly on the mid-points, a tick test is adopted as follows to classify the trade. If the current traded price is above (below) the previous price, the trade is an up-tick (down-tick), and is classified as an ask (a bid) trade. If the current price is equal to the previous price, the trade is classified according to the trade before that previous price. A zero up-tick (i.e., the previous trade is traded at an up-tick) is classified as an ask trade, and a zero down-tick (i.e., the previous trade is traded at a down tick) is classified as a bid trade. The process stops when there have been no changes in the traded price in the last two transactions, and the trade will not be included in the analysis. The time difference between the current trade and the oldest transaction (for the purpose of identification) is restricted to a maximum of five minutes.

Construction of bid and ask prices of index portfolio

The Hang Seng Index is a value-weighted index. The current index value is the ratio of the current total market value of the index stocks divided by the total market capitalization at the previous day's close, multiplied by the value of the index at the previous day's close. Applying the Hang Seng Index (HSI) methodology, the ask (or purchase) price of the index portfolio at time (interval) α on day t is given by:

$$S^a_{t_\alpha} = \sum_{i=1}^{33} W_{it} P^a_{it_\alpha} \; ; \tag{1}$$

where $W_{it} = (I_{t-1} / \sum_{i=1}^{33} N_{it} P_{it-1}) N_{it-1}$. W_{it} is the market value weight for security i on day t; $P^a_{it_\alpha}$ is the ask price for stock i at t_α; P_{it-1} is the closing price of stock i on day t - 1; N_i is the number of

[3] Trading in futures contracts has since been switched to an electronic trading platform, the Hong Kong Automated Trading System (HKATS).

outstanding shares for stock i; and I_{t-1} is the level of the closing value of the HSI on the previous trading day.

Similarly, the bid (or selling) price of the portfolio is given by:

$$S^b_{t_\alpha} = \sum_{i=1}^{33} W_{it} P^b_{it_\alpha} \qquad (2)$$

where $P^b_{it_\alpha}$ is the bid price for stock i at t_α.[4]

The bid or ask index is based on the firm bid and ask quotes displayed on the computerized open limit order book and immediately executable in the electronic trading environment. It is calculated only when all 33 bid or ask prices of the component stocks are available for the interval.[5]

This procedure automatically eliminates the non-trading price problem. There is also little non-synchronous price problem because the quotes are retrieved from the screen simultaneously.

Calculation of arbitrage basis

If the futures is overpriced at time α on day t, a potentially profitable strategy is to short a futures contract at F, and simultaneously buy the index (portfolio) at S^a with borrowed funds. The cash dividends from the index portfolio are placed on deposit. If the futures is underpriced, a potentially profitable strategy is to short the index basket and hedge the position by going long a futures contract. The arbitrageur in this case has to compensate the stock lenders for cash dividends.[6]

Controlling for trading costs C, the futures is overpriced if at a particular time α on day t, $F_{t_\alpha} > F^U_{t_\alpha}$, the bid futures price is above the corresponding upper no-arbitrage bound, where $F^U_{t_\alpha}$ is the upper no-arbitrage bound of the futures price; and

$$F^U_{t_\alpha} = S^a_{t_\alpha}(1+r)^{T-t} - \sum_{j=t}^{T-t-1} W_{it} D_{ij} (1+r_j)^{T-j} + C \qquad (3)$$

In equation (3) t and T (as fractions of a year) denote the initiation and the expiration day of an arbitrage portfolio; t_α denotes the particular time on day t; F is the futures price; and D_{ij} is the per share cash dividend for stock i at time j.

---

4 The contract multiplier is HK$50 per index point so that the actual number of shares for stock i in the index portfolio is equal to $50W_{it}$. Comparisons between the index and the futures price are based on the index value where W_{it} denotes the number of shares of stock i after conversion into index points. Similar considerations apply to the calculation of dividends for the index portfolio.

5 The first constructed bid or ask index is usually observed a few minutes after the morning session commences. This finding reflects the non-trading problem especially during the morning open.

6 We ignore the bid-ask price differential in the actual calculation because the futures were highly liquid, with a very low bid-ask spread.

By parallel reasoning, the futures is underpriced if $F_{t_\alpha} < F_{t_\alpha}^L$, where $F_{t_\alpha}^L$ is the lower no-arbitrage bound for the futures price:

$$F_{t_\alpha}^L = S_{t_\alpha}^b (1+r)^{T-t} - \sum_{j=t}^{T-t-1} W_{it} D_{ij} (1+r_j)^{T-j} - C \tag{4}$$

There is no arbitrage opportunity if $F_{t_\alpha} < F_{t_\alpha}^U$ or $F_{t_\alpha} > F_{t_\alpha}^L$.

I examine the mispricings relative to the corresponding upper and lower bounds of the futures price. Overpricing in the futures relative to the upper no-arbitrage bound is defined as:

$$e^+ = \frac{F_{t_\alpha} - F_{t_\alpha}^U}{F_{t_\alpha}^U} \quad for \quad F_{t_\alpha} > F_{t_\alpha}^U \tag{5}$$

Similarly, underpricing relative to the lower no-arbitrage bound is defined as:

$$e^- = \frac{F_{t_\alpha} - F_{t_\alpha}^L}{F_{t_\alpha}^L} \quad for \quad F_{t_\alpha} < F_{t_\alpha}^L \tag{6}$$

And use three series of the arbitrage basis. The basis assuming zero costs does not consider any trading costs except for the difference between the bid and ask index prices (i.e., C is set equal to zero in the no-arbitrage bound equations). The basis depending trading costs C factors in the various trading fees and charges for member firms. The basis depending on the average of the bid and ask index price serves as the benchmark.

Measures of order imbalance

Following Blume, MacKinlay, and Terker (1989), I assume the (dollar) order imbalance of an individual stock is equal to its dollar volume crossed at the ask price minus the dollar volume crossed at the bid price within a particular interval. As stock quotes are refreshed every 30 seconds, I first obtain the dollar order imbalance for each stock over each 30-second interval. Aggregate dollar order imbalance for the index within a 30-second time interval is obtained by summing the individual order imbalances of constituent stocks within the interval.

To reduce the scale problem in the order-imbalance measure, I convert the aggregate dollar order imbalance into an equivalent number of index baskets by dividing the measure by the product of HK$50 and the average of the reported bid and ask indices at the end of the 30-second time interval. This is the *order imbalance* (in index baskets) referred to hereafter.[7]

7 Results of tests using a dollar order imbalance measure are similar to those reported measuring order imbalance in index baskets.

4. Empirical results and interpretation

There are six different categories of arbitrage basis (or pricing error):

C1: The futures price is above the no-arbitrage upper bounds based on member cost assumptions.

C2: The futures price is between the no-arbitrage upper bounds based on member costs and zero trading cost assumptions.

C3: The futures price is between the no-arbitrage upper bounds based on zero cost trading cost assumptions and the fair futures price (F*).

C4: The futures price is between the fair futures price and the no-arbitrage lower bound based on zero trading cost assumptions.

C5: The futures price is between the no-arbitrage lower bounds based on member costs and zero trading cost assumptions.

C6: The futures price is below the no-arbitrage lower bounds based on member cost assumptions.

C1 (trading cost & bid/ask spread)
C2 (bid/ask spread only)
C3
C4
C5 (bid/ask spread only)
C6 (trading cost & bid/ask spread)

We expect order imbalance to be positive for categories C1 through C3 and negative for categories C4 through C6. We also expect order imbalance to be greater in C1 than in C2, and greater in C2 than in C3. Order imbalance in C6 is expected to be the most negative of C4, C5, and C6.

Table 1 summarizes the percentage basis pricing errors across the six categories.

Table 2 shows the summary statistics of order imbalance (in HK$) across the six categories. Pairwise comparisons show that the means of (dollar) order imbalance in category 1 and 2 are not significantly different. The means of all other consecutive pairs are significantly different at any reasonable confidence level (these results are not reported here to conserve space, but are available upon request). Hence, the results are largely consistent with the proposition that a positive basis is associated with positive order imbalance and a negative basis is associated with negative order imbalance. Moreover, large positive order imbalance is associated with a high positive basis, and large negative order imbalance is associated with a high negative basis.

Table 3 is similar to Table 2, except that it measures the order imbalance in number of index baskets. The results in Table 3 are about the same as those in Table 2.

Finally, I test the composite hypothesis that the order imbalance in category 1 is greater than the imbalance in 2, the imbalance in 2 is greater than in 3, and so on. Table 4 shows the Jonckheere and Terpstra nonparametric test results for ordered alternatives for both HK dollar and number of index basket measures of order imbalance. The results indicate rejection of the null hypothesis according to the dollar order imbalance but not the order imbalance in index baskets.[8]

4.1 Asymmetry in the arbitrage basis

Table 5 shows distributions of the three measures of arbitrage basis. For all three series, there is both greater and more frequent negative arbitrage basis than positive. These results reflect the inefficiencies and the additional cost associated with short-stock arbitrage when the futures is underpriced.

4.2 Order imbalance and arbitrage basis

The hypothesis in the relation between order imbalance and arbitrage basis is that the arbitrage basis should be positively related to order imbalance, and that negative order imbalance has more of an impact than positive order imbalance. We test these hypotheses with both interday and intraday data using the regression model:

$$e = \alpha_0 + \alpha_1 OI + \alpha_2 D(OI); \mathrm{D} = 0 \text{ if } \mathrm{OI} > 0; \mathrm{D} = 1 \text{ if } \mathrm{OI} < 0$$

e is expected to be positively related to order imbalance, and $\alpha_1 > 0$; moreover, if negative order imbalance has a greater impact than positive order imbalance, $\alpha_2 > 0$.

Findings measuring order imbalance in index baskets are reported; results using the dollar order imbalance are practically identical and are not reported here for reasons of space.

Panel 1 of Table 6 shows the results with daily data, and panel 2 shows results with intraday data sampled every minute. Both interday and intraday results show that negative order imbalance has a significant and greater impact than positive order imbalance on the arbitrage basis. The sign of the coefficient is positive and consistent with the proposition. Positive order imbalance has a significant impact on the basis on an intraday basis.

4.3 Intraday relationships: Sign of arbitrage basis and convergence time

Table 7 summarizes the time it takes for an observed arbitrage basis to disappear. Consistent with the findings on the asymmetry of the positive and negative basis, it takes more time on average for a negative basis to disappear than a positive basis. A pairwise t-test shows that the difference is significant for the arbitrage basis in the zero trading cost category.

---

8 Thanks to Kin Lam for suggesting this test.

4.4 Order imbalance and convergence time for a pricing error

Whether the convergence time is positively related to the extent of order imbalance and whether negative order imbalance has an asymmetrically greater impact than positive order imbalance is tested by the regression:

$$\text{Convergence time} = \delta_0 + \delta_1 |OI| + \delta_2 D |OI|$$

All three regression results in Table 8 show that the time it takes for an error to converge to zero is significantly related to the extent of the order imbalance. Moreover, the arbitrage basis series based on the mid-quote index and that based on the member cost assumption show that negative order imbalance has an asymmetrically greater effect on convergence time.

Differential impacts of positive versus negative observed pricing errors on the convergence time for positive and negative errors are tested by the regressions:

$$e^{+} : \text{Time} = \delta_0 + \delta_1 OI$$

$$e^{-} : \text{Time} = \tau_0 + \tau_1 OI$$

The results are reported in Table 9.

We expect $\delta_1 > 0$ if the convergence speed for positive errors is positively related to order imbalance. That is, positive order imbalance holds up positive pricing errors, while negative order imbalance speeds up convergence for positive errors. We expect $\tau_1 < 0$ if the convergence speed for negative errors is negatively related to order imbalance. That is, we expect negative order imbalance to delay convergence for negative errors, while positive order imbalance slows the convergence.

The results in Table 9 show that order imbalance does not significantly impact the correction time for positive errors but it does significantly affect the correction time for negative errors. Moreover, consistent with expectations, negative order imbalance delays error correction, while positive order imbalance speeds it up.

5. Conclusion

Complete transaction records of futures and index stocks, with bid/ask price quotes, let us examine how and whether the arbitrage basis between index futures and the underlying cash index is related to the direction and the extent of the aggregate order imbalance of the index stocks. The data cover an extended time period that includes the 1987 Asian financial crisis.

The results indicate that pricing errors are positively related to the aggregate order imbalance in the underlying index stocks; moreover, negative order imbalance has a stronger impact than positive order imbalance on the spread. The time it takes for a mispricing to converge to zero is positively related to the extent of order imbalance, and negative order imbalance has a greater impact than positive order imbalance. These results can be attributed to the additional cost and inefficiencies of short stock arbitrage.

These results support the conjecture that index arbitrageurs are important providers of liquidity in the futures market when the stock market moves away from equilibrium.

References

Bhatt, S., and Cakici, N. (1990), "Premiums on Stock Index Futures - Some Evidence," *Journal of Futures Markets*, 10: 367-75.

Blume, M.E., MacKinlay, A.C., and Terker, B. (1989), "Order Imbalances and Stock Price Movements on October 19 and 20, 1987," *Journal of Finance*, 44(4): 827-48.

Breen, W.J., Hodrick, L.S., and Korajczyk (2002), "Predicting equity liquidity," *Management Science*, 48(4): 470-83.

Chan, K. (1992), "A Further Analysis of the Lead Lag Relation between the Cash Market and Stock Index Futures Market," *Review of Financial Studies*, 5: 123-52.

Chan, K., Chung, Y.P., and Johnson, H. (1993), "Why Options Prices Lag Stock Prices: A Trading-based Explanation," *Journal of Finance*, 48(5): 1957-67.

Chan, K. and Fong, W.M. (2000), "Trade Size, Order Imbalance, and the Volatility - Volume Relation," *Journal of Financial Economics*, 57: 247-73.

Chen, N.F., Cuny, C.J., and Haugen R.A. (1995), "Stock Volatility and the Levels of the Basis and Open Interest in Futures Contracts," *Journal of Finance*, 50(1): 281-300.

Cornell, B., and French, K. (1983), "Taxes and the Prices of Stock Index Futures," *Journal of Finance*, 38: 675-94.

Draper, P., and Fung, J.K.W. (2003), "Discretionary government intervention and the mispricing of index futures," *Journal of Futures Markets*, 23(2): 1159-89.

Easley, D., O'Hara, M., and Srinivas, P.S. (1998), "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53(2): 431-65.

Fung, J. K.W. (2004), "Order Imbalance and the Pricing of Index Futures," Working Paper, Hong Kong Baptist University.

Fung, J.K.W., and Draper, P. (1999), "Index Arbitrage Opportunities and Short Sales Constraints," *Journal of Futures Markets*, 10(2): 695-715.

Fung, J.K.W., and Jiang, L. (1999), "Restrictions on Short-selling and Spot-Futures Dynamics," *Journal of Business, Finance and Accounting*, 26(1) and (2): 227-48.

Grossman, S.J. (1988), "An Analysis of the Implications for Stock and Futures Price Volatility of Program Trading and Dynamic Hedging Strategies," *Journal of Business*, 61(3): 275-98.

Harris, L. (1989), “The October 1987 S&P 500 Stock-Futures Basis,” *Journal of Finance*, 44(1): 77-99.

Harris, L., Sofianos, G. and Shapiro, J.E, (1994), “Program Trading and Intraday Volatility,” *Review of Financial Studies*, 7(4): 653-85.

Jiang, L., Fung, J.K.W., and Cheng, L.T.W. (2001), “Lead-lag Relationship between Spot and Futures Markets under Different Short-Selling Regimes,” *Financial Review*, 38: 63-88.

Kawaller, I.G. (1987), “Determining the Relevant Fair Value(s) of S&P 500 Futures: A Case Study Approach,” *Journal of Futures Markets*, 11(4): 453-60.

Kawaller, I.G., Koch, P.D. and Koch, T.W., (1987), “The Temporal Price Relationship Between S&P 500 Futures and the S&P Index,” *Journal of Finance*, 42(5):1309-29.

Kempf, A. (1998), “Short Selling, Unwinding, and Mispricing,” *Journal of Futures Markets*, 18(8): 903-23.

Kleidon, A.W. (1992), “Arbitrage, Nontrading, and Stale Prices: October 1987,” *Journal of Business*, 65(4): 483-507.

Kleidon, A.W., and Whaley, R.E. (1992), “One market? Stocks, futures, and options during October 1987,” *Journal of Finance*, 47(3): 851-77.

Lee, C.M.C., and Ready, M.J. (1991), “Inferring Trade Direction from Intraday Data,” *Journal of Finance*, 46(2): 733-46.

Locke, P.R., and Sayers, C.L. (1993), “Intra-Day Futures Price Volatility: Information Effects and Variance Persistence,” *Journal of Applied Econometrics*, 8: 15-30.

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?” *Journal of Finance*, 49(2): 479-513.

Modest, D.M., and Sundaresan, M. (1983), “The Relationship between Spot and Futures Markets,” *Journal of Futures Markets*, 3: 15-42.

Neal, R. (1996), “Direct Tests of Index Arbitrage Models,” *Journal of Financial and Quantitative Analysis*, 31(4): 541 - 62.

Roll, R., Schwartz, E., and Subrahmanyam, A. (2005), “Liquidity and the Law of One Price: The Case of the Futures/Cash Basis,” *Journal of Finance,* forthcoming.

Shleifer, A., and Vishny, R.W. (1997), “The Limits of Arbitrage,” *Journal of Finance*, 52(1), 35-55.

Sofianos, G. (1993), “Index Arbitrage Profitability,” *Journal of Derivatives*, 1: 6-20.

Stoll, H.R., and Whaley, R.E. (1990), “The Dynamics of Stock Index and Stock Index Futures Returns,” *Journal of Financial and Quantitative Analysis*, 25(4): 441-68.

Yadav, P.K., and Pope, P.F. (1990), “Stock Index Futures Arbitrage: International Evidence,” *Journal of Futures Markets*, 10(6): 573-603.

Yadav, P.K., and Pope, P.F. (1994), “Stock Index Futures Mispricing: Profit Opportunities or Risk Premia?” *Journal of Banking & Finance*, 18: 921-53.

Table 1. Summary statistics of the percentage basis (F - F*/F*) in basis points within the 6 categories of futures price positions

Category	N	Mean	t-stat	std	min	med	max	p95	p99
1	1227	93.49	166.56	19.66	72.00	87.74	196.85	138.91	168.51
2	9260	49.40	356.01	13.35	25.59	46.73	109.44	73.52	86.24
3	53484	14.22	335.43	9.81	0.00	12.79	59.10	31.96	39.82
4	77493	-18.10	-450.43	11.19	-62.28	-17.34	0.00	-1.76	-0.35
5	45461	-51.67	-808.12	13.63	-99.32	-50.32	-24.48	-31.76	-27.91
6	10739	-122.17	-159.91	79.18	-753.82	-102.13	-62.55	-74.94	-67.19

Category 1: The futures price is above the no-arbitrage upper bounds based on members' trading cost assumptions.

Category 2: The futures price is between the no-arbitrage upper bounds based on members' and zero trading cost assumptions.

Category 3: The futures price is between the no-arbitrage upper bounds based on zero cost trading cost assumptions and the fair futures price (F*).

Category 4: The futures price is between the fair futures price and the no-arbitrage lower bound based on zero trading cost assumptions.

Category 5: The futures price is between the no-arbitrage lower bounds based on zero cost and members' trading cost assumptions.

Category 6: The futures price is below the no-arbitrage lower bounds based on members' trading cost assumptions.

F* indicates the fair value of the futures based on the average of the concurrent bid and ask index value.

Table 2. Summary statistics of order imbalance (in HK dollar) corresponding to the 6 categories

Category	N	Mean	t-stat	std
1	1227	1312349	3.76	12242129
2	9260	988078	10.27	9262348
3	57131	247111	7.72	7652979
4	77493	-153834	-6.11	7006380
5	45461	-652441	-17.59	7908086
6	10739	-1706176	-18.14	9746344

Order imbalance in HK dollars for a particular 30-second time interval is equal to the aggregate of the dollar order imbalance for all index stocks in the interval. The t-statistics test the null hypothesis that the mean is equal to zero. Pairwise comparisons show that the means of (dollar) order imbalance in category 1 and 2 are not significantly different. The means of all other consecutive pairs are significantly different at any reasonable confidence level (results not reported here to conserve space).

Table 3. Summary statistics of order imbalance (in number of index baskets) corresponding to the 6 categories

Category	N	Mean	t-stat	std	min	med	max	p95	p99
1	1227	116.52	3.65	1117.35	-5361.79	0.00	14153.64	1950.86	3710.25
2	9260	89.11	10.68	803.19	-7242.79	5.21	13505.02	1378.89	2770.50
3	53484	19.69	7.20	632.91	-8478.04	-4.48	19736.84	896.75	2103.55
4	77493	-12.72	-6.35	557.36	-8263.97	-4.43	12690.79	720.54	1711.73
5	45461	-51.98	-17.62	629.20	-9218.59	-9.66	7755.05	753.21	1746.94
6	10739	-156.74	-17.93	905.92	-8673.15	-50.73	9032.48	1001.43	2217.15

Order imbalance in index basket for a particular 30-second time interval is equal to the aggregate (dollar) order imbalance of the interval divided by the value of the index basket based on the middle of the bid/ask index quotes at the end of the interval. The t-statistics correspond to the test on the null hypothesis that the mean is equal to zero. Pair-wise comparisons show that the means of order imbalance in category 1 and 2 are not significantly different. The means of all other consecutive pairs are significantly different at any reasonable confidence level. These results are not reported here to conserve space but are available upon request.

Table 4. Jonckheere and Terpstra nonparametric test for ordered alternatives

Test	I	II
J-stat	(19.14)***	-8.09

Test I: Order imbalance in HK$ across 6 categories (intraday).

Test II: Order imbalance in index basket across the 6 categories (intraday).

J-statistics allow us to reject the null hypothesis with order imbalance measured in HK$.

Table 5. Summary statistics of arbitrage basis under different trading cost assumptions for May 1, 1996 to April 30, 1998

		N	Mean	Std	Min	Med	Max	Skewness
Panel 1:	**e**	103994	-0.0018	0.0045	-0.0777	-0.0014	0.0241	-3.1001
Arbitrage	**e+**	35051	0.0022	0.0020	0.0000	0.0017	0.0241	1.9764
basis based	**e-**	68943	0.0038	0.0040	0.0000	0.0031	0.0777	6.0292
on the mid-index	**|e|**	103994	0.0033	0.0035	0.0000	0.0025	0.0777	6.1388
Panel 2:	**e**	30146	-0.0047	0.0095	-0.1430	-0.0034	0.0281	-6.1347
Arbitrage	**e+**	4383	0.0046	0.0037	0.0000	0.0035	0.0281	1.3446
basis assuming	**e-**	25763	0.0063	0.0092	0.0000	0.0041	0.1430	7.4735
zero trading cost	**|e|**	30146	0.0060	0.0087	0.0000	0.0040	0.1430	7.7903
Panel 3:	**e**	5104	-0.0077	0.0163	-0.1337	-0.0036	0.0191	-4.6180
Arbitrage	**e+**	467	0.0035	0.0028	0.0001	0.0027	0.0191	1.8823
basis assuming	**e-**	4637	0.0088	0.0167	0.0000	0.0041	0.1337	4.5840
member trading cost	**|e|**	5104	0.0083	0.0160	0.0000	0.0040	0.1337	4.8092

Two-sample t-test of the equality of the means of $e+$ and $e-$ allows us to reject at any reasonable confidence level that the mean magnitudes are equal. This implies that the negative basis is on average greater than the positive basis.

Table 6. Impact of order imbalance (in index baskets) on index-futures basis

Model: $e = \alpha_0 + \alpha_1 OI + \alpha_2 D(OI)$; D = 0 if OI > 0; D = 1 if OI < 0, D = 0 otherwise
H1: e is positively related to OI; $\alpha_1 > 0$
H2: negative OI has a greater impact than positive OI; $\alpha_2 > 0$

	Panel 1 Interday Regression of Errors	Panel 2 Intraday Regression of Errors (1 Min)
Dependent Variable	**(F - F*)/F*[1]**	
N	479	93820
Intercept	-0.21682 (-3.82)***	-0.00166 (-91.18)***
OI	0.00006692 (0.60)	0.00001027 (9.12)***
D(OI)	0.00051664 (3.11)***	0.00002566 (14.61)***
R2	0.0983	0.0134
F	25.99 0.00	638.73 0.00

[1](F - F*)/F* is the error in percentage of the fair value of the futures (F*). F* is defined by the middle of the concurrent bid and asked index quotes.

T values are reported in parentheses. *Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.

Panel 1: The dependent variable is the sum of the arbitrage basis in a trading day; OI is the aggregate order imbalance for the day. This table shows the result when the order imbalance in number of index baskets. The order imbalance in index basket for a particular 30 second time interval is equal to the order imbalance of the interval divided by the value of an index basket based on the middle of the bid/ask index quotes at the end of the interval. The aggregate daily order imbalance is equal to the sum of the order imbalance (in basket) defined every 30 second during the particular day.

Panel 2: The dependent variable is the arbitrage basis measured every 30-second; OI is the order imbalance in index basket for the corresponding 30 second time interval.

Table 7. Summary statistics of time for a mispricing to disappear (in seconds)

			N	Mean	Std
Basis (mid-index)	e	Time	6119	768.57	2112.46
	e+	Time	3083	750.63	2063.61
	e-	Time	3036	786.78	2161.13
Zero cost	e	Time	4805	431.38	1571.30
	e+	Time	1426	278.39	1119.29
	e-	Time	3379	495.93	1723.00
Member cost	e	Time	1607	299.55	1155.50
	e+	Time	250	257.76	939.92
	e-	Time	1357	307.25	1191.11

Time measures the duration for an initially observed pricing error to disappear. All errors observed within the time interval that are of the same sign as the initial error are omitted from the analysis. Two sample t-test of mean shows that the time for a negative mispricing to converge to zero is on average larger than that for the positive mispricing, for the zero cost category. Wilcoxin sign rank test on the medians of the convergence time show that the result is consistent with the mean test for both zero and member cost categories.

Table 8. Order imbalance of the time for a mispricing to disappear (in seconds)

Model: $Time = \beta_0 + \beta_1 |OI| + \beta_2 D|OI|$;
Time measures the duration for an initially observed pricing error to converge to zero. All errors observed within the time interval that are of the same sign as the initial error are omitted from the analysis. OI measures the total order imbalance within the time interval.

H1: Increase in the magnitude of OI delays convergence; $\beta_1 > 0$
H2: Negative OI induces greater delay against convergence; $\beta_2 > 0$

	N	Intercept	|OI|	D(|OI|)	R2	F
(F-F*)/F*[1]	5794	167.06367	9.92523	3.23792	0.4929	2814.50
		(7.66)***	(50.77)***	(11.35)***		0.00
Zero cost	4792	43.81211	5.54884	-0.24578	0.4828	2235.80
		(2.53)**	(42.45)***	(-1.55)		0.00
Member	1605	50.41961	3.00817	0.75618	0.3764	483.75
		(2.05)**	(10.02)***	(2.39)**		0.00

Table 9. Differential impacts of order imbalance on positive and negative arbitrage basis

Model 1: $e^+ : Time = \delta_0 + \delta_1 OI$

Model 2: $e^- : Time = \tau_0 + \tau_1 OI$

H1: Conditional on positive arbitrage basis, the convergence time is positively related to order imbalance; i.e., $\delta 1 > 0$

H2: Conditional on negative arbitrage basis, the convergence time is negatively related to order imbalance; $\tau 1 < 0$

		N	Intercept	OI	R2	F
(F-F*)/F*	e+	3082	750.79097	-0.02845	0.0000	0.00
			(20.07)***	(-0.04)		0.97
	e-	3035	736.17880	-5.94169	0.0227	70.47
			(18.76)***	(-8.39)***		0.00
Zero cost	e+	1425	198.03425	4.37285	0.0666	101.53
			(6.66)***	(10.08)***		0.00
	e-	3378	373.97795	-6.32818	0.0932	346.96
			(12.90)***	(-18.63)***		0.00
Member	e+	249	223.13680	1.49406	0.0107	2.68
			(3.55)***	(1.64)		0.10
	e-	1356	195.60880	-4.39556	0.0878	130.48
			(6.04)***	(-11.42)***		0.00

*Significant at 10% level. **Significant at 5% level. ***Significant at 1% level.