ELSEVIER

Journal of Empirical Finance 10 (2003) 305–320

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

How much do locals contribute to the price discovery process?

Kingsley Fong[a,*], Ralf Zurbruegg[b]

[a] School of Banking and Finance, University of New South Wales, Sydney, New South Wales NSW 2052, Australia
[b] School of Commerce, University of Adelaide, Australia

Abstract

This paper studies the relative contribution of floor brokers and local floor members (locals) towards the price discovery process. Using detailed Share Price Index (SPI) futures contract data from the Sydney Futures Exchange (SFE), we construct 1- and 5-min series of volume weighted average prices for both floor brokers and locals. Estimates of Gonzalo and Granger [J. Bus. Econ. Stat. 13 (1995) 27] plus Hasbrouck [J. Finance 50 (1995) 1175] information shares reveal that trades initiated by locals account for approximately one-third of all price discovery within the market. Also, it is found that the locals' share of price discovery is higher than their share of trading activity. However, when there is high volatility or when observation intervals are lengthened, locals' contribution to price discovery declines. We conclude that although floor brokers are the main contributors, locals do play a significant role within the price discovery process.

JEL classification: G14; G12
Keywords: Local-broker trading; Cointegration; Price discovery

1. Introduction

Price discovery is the dynamic process of finding an equilibrium price. Understanding the source of price changes and how external information is incorporated into prices enhances investors' confidence in a market and improves efficiency in the pooling of capital and sharing of risk. Quantification of the contribution by different trading

* Corresponding author. Tel.: +61-2-9385-4932; fax: +61-2-9385-6347.
E-mail address: K.Fong@unsw.edu.au (K. Fong).

0927-5398/02/$ - see front matter
doi:10.1016/S0927-5398(02)00053-1

mechanisms, markets, or even traders to the price discovery process can also help exchanges to design better markets, as well as provide useful information for investors to select where they might want to trade, e.g. to avoid trading against traders with superior information. This paper studies the relative contribution of different groups of traders on the Sydney Futures Exchange (SFE) trading floor towards the price discovery process.

There are two main types of traders in floor-based futures markets—local floor members (locals) and brokers. Locals are full-time principal traders, trading for themselves physically on the floor and observe information in real-time. Brokers trade as agent on behalf of other traders. Silber (1984) argues that floor specific information such as the order flow and the identity of traders are important sources of pricing information for locals[1]. For this reason, Massimb and Phelps (1994) suggest that floor-based systems with the assistance of locals provide superior liquidity and price discovery functions than alternative systems. Martens (1998) finds evidence of differential price discovery across trading systems. He compares price discovery in German Bund futures that are simultaneously traded in LIFFE and DTB and find that during high volatility periods, floor-based systems contribute more to price discovery than screen-based systems. The reverse pattern is observed during low volatility periods. Whether the result is due to local trading is not investigated.

In order to assess what impact locals can have upon the price discovery process, we utilise both Gonzalo and Granger (1995) as well as Hasbrouck's (1995) information share methodologies to measure the relative importance that brokers and locals have towards this process. Specifically, using the unique settlement and quote data of the Share Price Index futures contracts traded on the Sydney Futures Exchange, we find that floor brokers are the main contributors to the price discovery process in the futures trading floor. However, there is also strong evidence that locals aid the price discovery process by using floor specific information. In particular, local's share of price discovery is higher than their share of trading activity.

Our sample period straddles over the 1997 global stock market crash, which provides an opportunity to study the relationship between volatility and the locals' share of price discovery. By subdividing the sample, we compare locals' share of price discovery in a high volatility period and that in a low volatility period. We find that as floor specific information (relative to off-floor information) becomes less important in the high volatility period, the contribution to price discovery by locals reduces.

In addition to contributing to the price discovery literature, this paper also extends the limited research examining the behavior of locals[2]. Manaster and Mann (1996) investigate locals' trading in the Chicago Mercantile Exchange. Using trade-by-trade data, they find that locals behave more like short-term aggressive traders instead of passive market

[1] Forster and George (1992) also acknowledge the importance of information that can only be observed on trading floors. In the context of FX markets, Ito et al. (1998) name this type of trading (non-fundamental) information as semi-fundamental information.

[2] One reason for the small amount of research in this area is that few exchanges make available detailed transactions data that separates whether trades are carried out by either locals or brokers.

makers. Specifically, locals reduce their inventory by up to 50% every trade, plus trade at higher prices when their inventory is positive and at lower prices when their inventory is negative. The reason suggested for this is that locals have a slight information advantage in the short-run from on-the-floor trading. In addition, Frino and Jarnecic (2000) confirm the aggressiveness of locals' trading using both trade and quote data from the SFE. They find that locals initiate trades and pay the bid-ask spread only marginally less often than when they provide liquidity and receive the bid-ask spread. Finally, Fong (2001) models the order flow of brokers and locals and finds that local order flows lead broker order flows in 5-min intervals. He also concludes that locals have a significant short-term execution and information advantage over off-floor traders, explaining why they behave as short-term aggressive traders. None of these studies, however, address the issue of price discovery.

The rest of the paper is structured as follows. Section 2 describes the data and Section 3 provides the preliminary statistics. Section 4 details the econometric method while Section 5 reports the estimation results and Section 6 presents a robustness test. A conclusion is then provided in Section 7.

2. Data

The Share Price Index (SPI) futures contract traded on the Sydney Futures Exchange (SFE) was the first index futures contract listed outside the U.S. Prior to the SFE's automation in December 1999, SPI futures were traded on the floor by open outcry between 9:50 and 12:30 in the morning, and 14:00 and 16:15 in the afternoon. Overnight trading took place via their electronic limit order book, SYCOM.

The benefit of examining trading behaviour at the SFE is that it maintains detailed records of trades that are not available from other exchanges. Specifically, this paper utilises confidential settlement and quotation data for the SPI futures contract during daytime trading hours from July 24th 1997 to June 25th 1998. This data was provided to us by the Sydney Futures Exchange via the Securities Industry Research Centre of Asia-Pacific (SIRCA).

The settlement data contains the time of trade to the minute, price, number of contracts traded, and most importantly, whether the counterparties are trading for a brokerage house, i.e. brokers, or as a local member, i.e. locals. Also, we use the nearby contract with a rollover date set at two trading days prior to expiry[3]. The settlement data is then matched with the quotation data. Finally, since the objective of this study is to measure the contribution by the two groups of traders to the price discovery process, we need to construct two price series: one for the brokers and one for the locals.

One way to construct the price series is to focus on the counterparties. Theoretically, there are three possible series: broker only, broker-local, and local only. Fong (2001) reports that local only trades account for only 3% of the number of trades for SPI futures.

[3] Unlike futures contracts trading in the U.S., the futures contract closest to expiry on the SFE remains the most liquid contract until the last 1–3 days to expiry.

Since infrequent trading poses statistical problems of spurious lagged correlations, which could lead to spurious conclusions, it is desirable to group this type of trades as broker-local trades. This grouping procedure results in two groups of trades, broker-only and broker-local. We call this method the counterparty method.

Another dimension of the trade data is the trade initiator. The trade initiator is the aggressive party in a trade, the market order user, and the consumer of liquidity. The passive party of a trade is the limit order user and liquidity provider. For studying price discovery, there is a strong case to focus on the aggressive party because he is the trader who causes a trade and the subsequent quote revision. However, the aggressor of a trade is not directly observed in the data and must be inferred. In order to do this, we compare the price of a trade to the matched quotation. If the price is at or higher than the ask quote, we say this is a buyer-initiated trade. If the price is at or lower than the bid quote, we say this is a seller-initiated trade. A trade is classified as either broker-initiated or local-initiated depending on whether the trade initiator is a broker or a local. We call this classification method the initiator method. Since this method generates richer and more meaningful series for studying price discovery, we will focus on this method in the following analysis and discussions. However, a comparison of results will also be made later using the broker-only and broker-local series.

3. Preliminary statistics

Table 1 contains the number of trades and contracts traded across different trader groups. Panel A shows that brokers initiate 75.6% of the trades and 80.1% of contracts traded. Frino and Jarnecic (2000) report that locals initiate trades slightly less than 50% of the time when they trade. This explains the difference between trading activity statistics across the two trade classification methods. Using the counterparty method, broker-only trades account for 48.1% of the number of trades (59% of contracts) while broker-local trades account for the remaining 51.9%. Notice that 48.1% plus one half of 51.9% is close to 75.6%. The counterparty classification result is broadly consistent with findings from the Chicago Mercantile Exchange reported in Manaster and Mann (1996).

One problem in defining the price series is that trades occur at irregular time intervals. In addition, while some trades on the trading floor could take place simultaneously, most are not. In order to study price discovery, we need to aggregate prices of the two trading groups along the same time scale. However, since these two groups of traders are trading the same securities on the same trading floor, aggregating prices over sufficiently long time intervals will eventually lead to identical price series for both types of traders, and hence the inability to find out the source of price discovery. Therefore, as a trade-off and also for comparison, we aggregate and examine the trade data over two intervals, 1- and 5-min. The lower frequency data series benefit from minimizing nonsynchronous trading, while the higher frequency data series will highlight the differences between the share of price discovery by locals and brokers.

Table 1
Trading volume and frequency

	Initiator method		Counterparty method	
	Broker-initiated	Local-initiated	Broker only	Broker-local
Panel A. Full sample				
Number of contracts traded				
Group total	1,691,306	419,066	1,245,723	865,418
Group percentage	80.1%	19.9%	59.0%	41.0%
Number of trades				
Group total	293,800	94,766	186,876	201,927
Group percentage	75.6%	24.4%	48.1%	51.9%
At 1-min intervals				
Average number of trades	4.6	1.5	2.9	3.2
Median number of trades	3	1	2	2
At 5-min intervals				
Average number of trades	22.8	7.4	14.5	15.7
Median number of trades	18	6	11	13
Panel B. High volatility sub-sample				
Number of contracts traded				
Group total	435,821	101,512	330,272	207,331
Group percentage	81.1%	18.9%	61.4%	38.6%
Number of trades				
Group total	73,669	23,434	47,466	49,696
Group percentage	75.9%	24.1%	48.9%	51.1%
Panel C. Low volatility sub-sample				
Number of contracts traded				
Group total	421,054	104,499	304,808	220,898
Group percentage	80.1%	19.9%	58.0%	42.0%
Number of trades				
Group total	79,515	25,931	48,919	56,587
Group percentage	75.4%	24.6%	46.4%	53.6%

Panel A also lists the average number of trades of the different trader groups in the two time scales. There are 4.6 trades per 1-min interval that are broker-initiated, with a median of 3 trades. The corresponding numbers for local-initiated trades are 1.5 and 1, respectively. Trade frequencies of broker-only and broker-local series are more evenly distributed, both have the mean number of trades of approximately 3 and medians of 2. At 5-min intervals, the values are five times that of 1-min intervals. Given that price discovery is inferred from the lead-lag structure between the price series, the low number of local trades at 1-min intervals (hence higher frequency on non-trading) introduces a potential bias to find a higher proportion of price discovery attributed to brokers at this frequency than at the 5-min interval.

Table 2 provides the descriptive statistics on return volatility during the full sample period and subperiods. Geometric compounding returns are computed using volume weighted average prices across all trades of a trader group within an interval. For brevity, we only report the broker-initiated and local-initiated returns statistics. Panel A contains

Table 2
Descriptive statistics

	Time Period	# Obs.	Series	Mean $\times 10^6$	Variance $\times 10^6$	Maximum	Minimum	Skewness	Excess kurtosis
PANEL A. 1-min frequency									
Full sample	24 July 97–	63,337	Local	0.476	0.558	0.020	− 0.017	0.638	56.636
	25 June 98		Broker	0.476	0.473	0.020	− 0.018	0.213	53.084
High volatility	1 Oct. 97–	15,798	Local	3.231	0.990	0.020	− 0.017	1.091	53.544
period	2 Jan. 98		Broker	3.226	0.852	0.020	− 0.018	0.396	53.260
Low volatility	2 Jan. 98–	17,976	Local	− 0.335	0.445	0.011	− 0.010	− 0.059	26.843
period	1 April 98		Broker	− 0.378	0.391	0.008	− 0.008	− 0.084	18.209
PANEL B. 5-min frequency									
Full sample	24 July 97–	12,621	Local	1.050	0.984	0.016	− 0.010	0.497	14.200
	25 June 98		Broker	1.050	0.094	0.022	− 0.011	0.951	27.655
High volatility	1 Oct. 97–	3152	Local	14.842	1.662	0.016	− 0.010	0.960	16.494
period	2 Jan. 98		Broker	14.758	1.638	0.022	− 0.011	1.722	33.651
Low volatility	2 Jan. 98–	3580	Local	− 1.640	0.087	0.005	− 0.007	− 0.236	2.729
period	1 April 98		Broker	− 1.220	0.085	0.005	− 0.006	− 0.251	2.501

Figures above are from logarithmic returns.

the information at 1-min intervals. There are 63,337 observations in the full sample. Since both locals and brokers are trading on the same trading floor, it is not surprising that the mean returns of the two series are very similar. The variance of returns for locals are slightly higher than that of brokers, possibly due to their lower trading frequency.

Panel A also shows the statistics for the high volatility subperiod (1st October 1997 to 2nd January 1998) that covers the 1997 global stock market crash and the low volatility subperiod (3rd January 1998 to 1st April 1998). During the high volatility subperiod, mean returns are much more positive than during the low volatility subperiod[4]. Also, the variance of returns during the high volatility subperiod is more than twice as high as that during the low volatility subperiod. As a visual display, Fig. 1 shows 1-min prices for the entire sample. It is evident from this that much of the price variation over the entire period occurred within the high volatility subperiod[5].

Given recent research pointing to locals being short-term traders[6] that rely on floor specific information, we expect their share of price discovery to be smaller in the high volatility subperiod. The reason for this is that the volatility in the high volatility subperiod is primarily due to an external factor, the global market crash. As it would be reasonable to assume that there is an abnormally large flow of information during this period from the outside world to the trading floor, floor specific information will not be as useful. However, it is interesting to note in panels B and C of Table 1 that there is no major

[4] The market crash took place overnight, while the next day, re-bounce took place during floor trading hours. Consequently, average return over floor trading hours is positive.

[5] The two subperiods chosen are representative of the results obtained from examining other similar periods, right up until the floor-based system concluded at the start of the year 2000.

[6] See Manaster and Mann (1996), Frino and Jarnecic (2000), and Fong (2001).

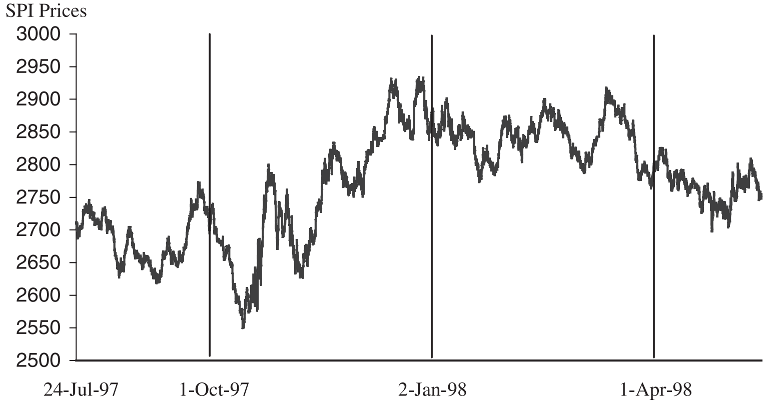

Fig. 1. SPI price series.

difference in locals' share of the number of trades and the number of contracts traded across the high and low volatility periods.

4. Methodology

Examining the price discovery process within financial markets has become a major focus of much recent empirical research. In particular, using similar econometric techniques to this paper, Harris et al. (1995) and Hasbrouck (1995) studied price discovery across different U.S. equity markets trading identical securities. They find that most of the price discovery occurs in the largest market, the New York Stock Exchange. This being most likely due to the relatively high trading volume and frequency of transactions on the NYSE[7].

The primary tools utilised in the above papers to examine price discovery tend to focus on the long- and short-run relationships asset prices have with each other. One simple method to study price discovery across multiple series is to examine Granger causality. Table 3 contains the results of the tests showing that there is bi-directional causality[8] of the

[7] Biais et al. (1999), Cao et al. (2000), and Madhavan and Panchapagesan (2000) also study price discovery in the pre-opening period in the Paris Bourse, NASDAQ, and NYSE, respectively. They find that traders signal with, and learn from, limit orders posted prior to the market opening, and that dedicated dealers facilitate price discovery relative to a pure computerized auction system. Also, Battalio et al. (1997) study the effect of SOES trades on volatility and price discovery. They find that SOES trades enhance pricing efficiency, instead of causing excess volatility, as they increase minute-by-minute volatility but dampen longer horizon volatility.

[8] All series are examined in first differences (returns) as Phillips and Perron (1988) unit root tests tabulated in Table 4 reveal all series having a single unit root. The Phillips–Perron methodology is chosen as it is consistent under conditions of heteroskedasticity and autocorrelation of unknown form.

Table 3
Standard Granger causality tests

	1-min frequency	5-min frequency
Full sample		
LOCAL returns do not cause BROKER returns	100.522	58.476
BROKER returns do not cause LOCAL returns	265.759	239.195
High volatility period		
LOCAL returns do not cause BROKER returns	28.577	11.361
BROKER returns do not cause LOCAL returns	70.633	58.869
Low volatility period		
LOCAL returns do not cause BROKER returns	32.173	22.407
BROKER returns do not cause LOCAL returns	77.286	65.199

All figures are F-statistics and have a significance level >0.0001.
Thirty-seven and eight lags are used to minimise the AIC for the 1- and 5-min series, respectively.

broker-initiated and local-initiated return series[9], although the magnitude of the values suggests it is in favour of brokers to locals. Moreover, the 5-min frequency data series extenuates the differences between the series, as in absolute values the locals' F-statistics (as well as in percentage change) diminish further. This preliminary result is consistent with our priori assumption that locals are likely to play a larger role in price discovery when measured over shorter time intervals.

The above Granger causality tests, however, cannot provide an accurate measure of the price discovery process for SPI index futures. One problem arises from Lindley's (1957) paradox where the use of large sample sizes can lead to the classical tests for significance to show significant results for insignificant economic values. This is demonstrated from the highly significant F-statistics that show bi-directional causal relationships between local and broker returns. The extent to which both participants influence each other is less clear.

Moreover, under the classical Granger-causality tests, no consideration is made for the long-run equilibrium between local and broker prices. As both types of traders are pricing the same instrument, the value of the long-run price for the asset must be the same. In fact, there must be a long-run implicit efficient price. This is shown in Table 4 where both Johansen (1991, 1995) maximum eigenvalue and trace tests show there is a single cointegrating relationship between the two series, indicating a single stochastic process (see Stock and Watson, 1988). From this fact, it is possible to separate the short- and long-run components of the series. The result of doing so will lead to the ability to distinguish whether locals or brokers are more influential in determining the long-run efficient price of the asset. In fact, if it is shown that one type of trader has no impact upon the implicit price of the assets, then we may say that it cannot have any value within the price discovery process. Either they are completely uninformed investors trading on random noise or merely follow other participants with a lag.

[9] We shall abbreviate these two series as broker-returns and local-returns for the remainder of this paper. Unless specified otherwise, the series we use are the broker-initiated and local initiated series.

Table 4
Unit root and cointegration tests

Phillips–Perron unit root tests	1-min frequency		5-min frequency	
	Local	Broker	Local	Broker
Log prices	– 1.900	– 1.885	– 2.051	– 2.052
First differences (returns)	– 341.008[a]	– 340.476[a]	– 103.575[a]	– 97.243[a]

Cointegration		1-min frequency	5-min frequency
Trace test	$r=0$	1244.830[a]	1162.073[a]
	$r \leq 1$	0.066	0.016
Maximum eigenvalue test	$r=0$	1243.551[a]	1162.057[a]
	$r \leq 1$	0.065	0.016

The Phillips–Perron tests above are conducted on logarithmic values for both series where the test statistic is corrected for heteroskedasticity and autocorrelation of unknown form using the Newey–West consistent estimate with a time trend. The number of truncation lags was set to 16 and 11, based on a floor function from the Newey–West estimator and the number of observations from the 1- and 5-min samples, respectively. Johansen cointegration statistics, for the hypothesis of either a rank of zero or a rank of no more than one, incorporates a 37- and 8-period lag structure determined by minimising the Akaike Information Criterion (AIC) for the 1- and 5-min samples, respectively. Critical values are obtained from Osterwald-Lenum (1992).

[a] Indicates rejection of the null hypothesis at the 1% significance level.

In order to examine this price discovery process more carefully, an error-correction model can be built to encompass the cointegrated nature of the series. Specifically, returns for locals, $r_{L,t}$, and brokers, $r_{B,t}$, can be expressed as:

$$r_{L,t} = a_L + \sum_{i=1}^{k} b_{L,i} r_{L,t-i} + \sum_{i=1}^{k} C_{B,i} r_{B,t-i} + \alpha_L (p_{L,t-1} - \beta_L p_{B,t-1}) \tag{1}$$

$$r_{B,t} = a_B + \sum_{i=1}^{k} b_{B,i} r_{B,t-i} + \sum_{i=1}^{k} C_{L,i} r_{L,t-i} + \alpha_B (p_{B,t-1} - \beta_B p_{L,t-1}) \tag{2}$$

where $p_{B,t}$ and $p_{L,t}$ are broker and local prices, and α is the adjustment coefficient for the error-correction term with β being the cointegrating vector. As with the standard Granger causality tests, b_i and c_i represent the lagged return coefficients up to k lags.

Moreover, rather than conducting causality tests on the lagged coefficients, and for the purposes of identifying the contribution each participant makes towards the price discovery process, focus can be placed upon both the permanent and transitory components of the model. Specifically, Gonzalo and Granger (1995) utilise the concept that a vector of market prices, p_t, can be expressed as a sum of these components:

$$p_t = f_t i_2 + g_t \tag{3}$$

where g_t is the (2×1) transitory component, f_t is a scalar representing the common factor, and i_2 is a (2×1) unit vector for the two types of traders. Gonzalo and Granger suggest that the common permanent component $f_t i_2$ can be identified by two conditions. The first is that the common factor is a weighted average of current market prices. The second is that the transitory component g_t has no long-run impact upon prices p_t. The first condition implies Eq. (4), where Θ is a (1×2) matrix of weights:

$$f_t = \Theta p_t \tag{4}$$

The second condition identifies Θ, as Gonzalo and Granger observe that there is only one weighting vector, Θ, for which g_t has no long-term impact on p_t. This is defined when $\Theta = \alpha^*$, where $\alpha^{*\prime}\alpha = 0$. That is to say that the weighting vector Θ can be estimated as a basis for the null space of the speed of adjustment coefficient vector, α. Once the elements of α^* have been normalized, they comprise a useful measure of the contribution of each market to the common permanent factor and thereby to price discovery.

5. Empirical results

Table 5 presents the estimates of Gonzalo–Granger (G–G) information shares. The first row shows the full sample estimates of the proportion of price discovery attributed to locals (Local weighting) and brokers (Broker weighting) at 1- and 5-min intervals. The results show brokers are unambiguously the main contributor to price discovery, accounting for 69.91% at 1-min intervals. Locals' contribution makes up the remaining 30.09%. Also, the shift in these ratios as the sampling frequency decreases to 5-min strongly supports the notion that locals contribute less to price discovery over longer time horizons. At 5-min intervals, their share drops to 25.62%. However, it should be emphasized that these estimates of locals' share of price discovery are higher than their share of trading activity that they initiated, which is 19.9% by contracts traded and 24.4% by number of trades.

The higher contribution of locals to price discovery in the shorter horizon than the longer 5-min horizon is consistent with the current research indicating that locals use short-lived floor specific information. As the frequency of observations declines, the usefulness of short-lived, floor-specific information is reduced[10] and thereby limiting the measured share of price discovery by locals.

It is also noteworthy to mention that the results of a smaller contribution by locals as the sampling interval moves from 1- to 5-min periods is despite a bias working in the opposite direction. Earlier, it was noted that with the relatively lower frequency of local-initiated trades at 1-min intervals, there would be a possibility of underestimating the locals' true contribution to price discovery. This bias is reduced when returns are computed over 5-min intervals because there are fewer observations of locals trading with stale prices, hence local returns do not appear as lagged behind broker prices as in the case of 1-min returns. Given this fact, it is important to stress that this bias and the shift in its effect across the two sampling frequencies should lead to a higher value for the locals' contribution to price formation, rather than less.

[10] For instance, floor specific information can help locals to predict order flow in the next minute. At 1-min intervals, the floor specific information spans the whole of the next observation interval. At 5-min intervals, it only spans one-fifth.

Table 5
Gonzalo–Granger information shares

	Time period	1-min frequency		5-min frequency	
		Local weighting	Broker weighting	Local weighting	Broker weighting
Full sample	24-Jul.-97–25-Jun.-98	30.09	69.91	25.62	74.38
High volatility period	1-Oct.-97–2-Jan.-98	25.71	74.29	8.75	91.25
Low volatility period	2-Jan.-98–1-Apr.-98	39.32	60.68	48.17	51.83

We now turn to price discovery in high and low volatility periods. The results suggest that locals' contribution to price discovery is lower during the high volatility period than during the low volatility period. Table 5 shows that, at 1-min intervals, locals' share of price discovery is 25.71% during the high volatility period and 39.32% during the low volatility period. At 5-min intervals, the contrast is even more dramatic, 8.75% and 48.17%, respectively. This is despite the fact that the share of trading volume across the high and low volatility periods are very similar. This pattern of results further reinforces the idea that locals trade on short-term floor specific information. As floor specific information becomes less important and off-floor information becomes more important, the contribution to price discovery by locals reduces[11].

For robustness, we also estimate the G–G information share model using the alternative series that only take into account counterparties identity. The estimates are tabulated in Table 6. One clear difference between Tables 5 and 6 is that estimates of locals' share of price discovery are higher when we use the brokers-only and broker-local series (Table 6), instead of the more meaningful classification that also incorporates trade initiator information (Table 5). This increase in the broker-locals' share is somewhat expected since more trades are now classified as local trades. Recall from Table 1 that local-initiated trades account for only 24.4% of all trades initiated and broker-local trades account for 51.9%. In fact, approximately half the broker-local trades are actually initiated by brokers rather than locals. As the trading frequency is more balanced between the two trade groups, the information shares are also likely to be more even. Furthermore, with an increased frequency of trades classified as "locals", there is a further reduction in the downward bias for the contribution locals make towards the price discovery process.

Despite the increase in estimates of the locals' share of price discovery using the counterparties classification, estimates in Table 6 still support our earlier results. First, the locals' share of price discovery is lower over longer horizons. Second, locals' share of price discovery is lower during the high volatility period than during the low volatility period. Third, whatever the sample or frequency used, locals do contribute to at least one

[11] The hypothesis of a lower locals' share of price discovery over longer horizons is clearly illustrated in the high volatility sample. However, it is not supported by the low volatility period estimates, although it can be argued that this is a result of the weakening of the bias for underestimating the locals' contribution on a 5-min frequency interval. As mentioned previously, the bias from using 1-min intervals could cause an estimated increase in the locals' share of price discovery when using 5-min prices, even if there is no real change.

Table 6
Alternative series Gonzalo–Granger information shares

	Time period	1-min frequency		5-min frequency	
		Local weighting	Broker weighting	Local weighting	Broker weighting
Full sample	24-Jul.-97–25-Jun.-98	47.89	52.11	35.13	64.87
High volatility period	1-Oct.-97–2-Jan.-98	42.49	57.51	23.45	76.55
Low volatility period	2-Jan.-98	56.38	43.62	51.37	48.63

quarter of the information shares, indicating their significance for price formation in this market.

6. Alternative methodology

So far, the G–G methodology has been utilised to identify the contribution locals and brokers make towards the common implicit price. However, an alternative estimation process also exists which theoretically incorporates a larger information set to determine the contribution each trader group makes towards the common price. Specifically, the G–G method identifies the common factor as a weighted average of current prices. However, this will exclude any information content held in past prices and requires a narrower definition of price discovery than is offered by Hasbrouck's (1995) technique. With Hasbrouck (1995), information shares arise from a random-walk decomposition with limited identification restrictions. This allows for the assumption that traders can update conditional expectations as new information enters the market. Information shares are therefore based upon reactions to price shocks, with each participant's contribution to the price discovery process being defined as the variation in efficient price innovations attributable to each type of trader. The description of the Hasbrouck procedure is provided in Appendix A, which is based on inverting the error-correction model detailed in Eqs. (1) and (2), before isolating the specific innovations attributable to either brokers or locals.

One disadvantage of the Hasbrouck methodology over G–G is that information shares attributable to each type of trader can only be determined within bounds due to the lack of identification of the parameters when there is cross-equation correlation. Furthermore, the size of the lower and upper bounds are also dependent upon the correlation between the innovations. For example, as the frequency of observations is lowered from 1- to 5-min, the correlations will be closer to one and identification becomes more difficult, leading to larger bounds. Therefore, the results tabulated in Table 7 show the upper and lower information bounds for the 1-min frequency observations only. The 5-min results provide too large boundaries to be useful for this study[12].

[12] Information bounds for both locals and traders were generally between 8% and 99% for the 5-min frequency data series. The sudden widening of the information bounds can be explained by examining the correlation coefficients. For the 1-min series, the correlation between broker and local returns is 0.14. On a 5-min basis, it is equal to 0.77.

Table 7
Hasbrouck information bounds

	Time period	Local weighting	Broker weighting
Full sample	24-Jul.-97 – 25-Jun.-98	(11.75, 41.90) 26.83	(58.10, 88.25) 73.18
High volatility period	1-Oct.-97 – 2-Jan.-98	(8.00, 38.38) 23.19	(61.62, 92.00) 76.81
Low volatility period	2-Jan.-98 – 1-Apr.-98	(20.96, 52.08) 36.52	(47.92, 79.04) 63.48

Table 7 also shows the mean information share for the upper and lower bounds computed for each trader group using the broker-initiated and local-initiated series. Using Hasbrouck's (1995) method[13], estimated locals' share of price discovery is 26.83% for the full sample period, 23.19% during the high volatility period, and 36.52% during the low volatility period. These estimates are consistent with the Gonzalo and Granger (1995) estimates reported in Table 5. Although the percentages are not exactly the same, the G–G information shares obtained do fall within the Hasbrouck bounds and differ by no more than 13%[14]. As a simple robustness test, the G–G results are relatively consistent with that of Hasbrouck's technique, despite the underlying differences between the two methodologies. These results confirm that locals do play a role in the price discovery process, and locals' contribution to price discovery is negatively related to market volatility originating from outside the exchange.

7. Conclusion

Using Share Price Index futures contract data from the Sydney Futures Exchange, we constructed 1- and 5-min series of volume weighted average prices for floor brokers and local floor members (locals). Estimates of the Gonzalo and Granger (1995) information shares show that at 1-min intervals, trades initiated by locals account for approximately 30% of the price discovery process in the market, while trades initiated by brokers account for the remaining 70%. At 5-min intervals, the contribution to price discovery by locals is still significant at approximately 26%.

Our sample period also includes the global stock market crash of October 1997 which allows for an interesting test on how local and broker contributions to price discovery vary when we split the data into high and low volatility periods. We find that locals' contribution to price discovery is much lower in the high volatility period than in the low volatility period. For the high volatility period, locals' contribution to price discovery is 25.71% at 1-min intervals. In the low volatility period, their contribution at 1-min intervals was 39.32%. For robustness, we also estimate the relative contribution of brokers and locals to price discovery based on Hasbrouck's (1995) methodology. The results are consistent with those of Gonzalo and Granger (1995), with the G–G results falling within Hasbrouck's information bounds.

[13] As well as having 37 lagged returns for the error correction model, the moving average representation contained 69 lags.

[14] The largest differences between the results of the two methodologies exist for the full sample period rather than the subsamples.

We conclude that floor brokers are the main contributors to the price discovery process in the futures trading floor. However, locals do aid the price discovery process in a significant way, most likely through the use of floor specific information. It is also extremely important to point out that their contribution to price discovery is actually higher than their share of trades initiated by them. However, as floor specific information becomes less important as the sample time horizon or volatility (off-floor information) increases, the contribution to price discovery by locals is reduced.

One possible implication from the above results is the cost of switching from open outcry to screen-based systems. Electronic trading may well enhance the price discovery process by increasing more active participation from off-floor traders who previously suffered an information disadvantage, or from informed traders who now benefit from anonymity. However, given that this paper has shown that locals contribute to more than their share of trading towards the price discovery process, their exclusion may lead to the removal of a possible valuable source of information.

Acknowledgements

We are grateful to Alex Frino at SIRCA and the Sydney Futures Exchange for the permission to use the data for this study. We also thank Theo Vermaelen and the anonymous referee for their encouragement and constructive comments.

Appendix A

Beginning with a standard error-correction model as in Eqs. (1) and (2), it is inverted into a vector moving average (VMA) representation such as that in Eq. (A.1).

$$\Delta p_t = \sum_{k=0}^{r} \psi(k) e_{t-k}, \tag{A.1}$$

where r is the number of moving average lags and

$$\psi(k) = \begin{bmatrix} \psi_{11}(k) & \ldots & \psi_{1n}(k) \\ \cdot & & \cdot \\ \psi_{n0}(k) & \ldots & \psi_{nn}(k) \end{bmatrix}, \tag{A.2}$$

The innovation vector e_t is presumed to have a zero mean, be serially uncorrelated, and have covariance matrix Ω. Each coefficient matrix $\psi(k)$ is an $(n \times n)$ matrix, where entry $\psi_{ji}(k)$ represents the impact of the innovation in the ith series on the price of the jth series in the $(t-k)$th period. The sum of the VMA coefficients can be defined as:

$$\psi(\mathrm{Sum}) = \sum_{k=0}^{q} \psi(k) = \begin{bmatrix} \sum_{k=0}^{q} \psi_{11}(k) & \dots & \sum_{k=0}^{q} \psi_{1n}(k) \\ \cdot & & \cdot \\ \sum_{k=0}^{q} \psi_{n1}(k) & \dots & \sum_{k=0}^{q} \psi_{nn}(k) \end{bmatrix}, \tag{A.3}$$

Hasbrouck proposes that the row vectors in ψ(Sum) should be equal if local and broker prices are cointegrated, as any disturbances to prices in the long-run must be the same. Allowing ψ to represent this common row vector and combining ψ with an innovation vector from a given period t produces ψe_t, which is the total long-run impact of the innovation in period t. Hasbrouck suggests that the sources of variance in this term are a measure of the contribution of each series to price discovery. The variance of ψe_t is presented as:

$$\mathrm{Var}(\psi e_t) = \psi \Omega \psi', \tag{A.4}$$

As is likely to be the case that the innovations in local and broker prices are correlated, a decomposition of the variances is required. Hasbrouck suggests that by triangularising the covariance matrix, upper and lower information share bounds can be computed. The triangularisation is performed by obtaining F, the Cholesky factorisation of Ω (i.e. we find F such that $\Omega = FF'$), where F is the lower triangular. The variance attributable to a particular market j is then $(\psi F)_j^2$ and the information share of trader type j is:

$$\mathrm{IS}_j = \frac{(\psi F)_j^2}{\psi \Omega \psi'}, \tag{A.5}$$

Hasbrouck notes that such a factorisation maximises the information share of the first price and so one must calculate upper and lower bounds for each information share by changing the ordering of the series in ψ and Ω.

References

Battalio, R.H., Hatch, B., Jennings, R., 1997. SOES trading and market volatility. Journal of Financial and Quantitative Analysis 32, 225–238.

Biais, B., Hillion, P., Spatt, C., 1999. Price discovery and learning during the preopening period in the Paris bourse. Journal of Political Economy 107, 1218–1248.

Cao, C., Ghysels, E., Hatheway, F., 2000. Price discovery without trading: evidence from the Nasdaq preopening. Journal of Finance 55, 1339–1365.

Fong, K., 2001. Market transparency and trading strategies. PhD Dissertation, University of Sydney.

Forster, M., George, T.J., 1992. Anonymity in securities markets. Journal of Financial Intermediation 2, 168–206.

Frino, A., Jarnecic, E., 2000. An empirical analysis of the supply of liquidity by locals in futures markets: evidence from the Sydney Futures Exchange. Pacific-Basin Finance Journal 8, 443–456.
Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. Journal of Business and Economic Statistics 13 (1), 27–35.
Harris, F.H., McInish, T.H., Shoesmith, G.L., Wood, R.A., 1995. Cointegration, error correction, and price discovery on informationally linked security markets. Journal of Financial and Quantitative Analysis 30, 563–579.
Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. Journal of Finance 50, 1175–1199.
Ito, T., Lyons, R., Melvin, M.T., 1998. Is there private information in the FX market? The Tokyo experiment. Journal of Finance 53, 1111–1130.
Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580.
Johansen, S., 1995. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford Univ. Press, Oxford.
Lindley, D.V., 1957. A statistical paradox. Biometrika 44, 187–192.
Madhavan, A., Panchapagesan, V., 2000. Price discovery in auction markets: a look inside the black box. Review of Financial Studies 13, 627–658.
Manaster, S., Mann, S., 1996. Life in the pits: competitive market making and inventory control. Review of Financial Studies 9 (3), 953–975.
Martens, M., 1998. Price discovery in high and low volatility periods: open outcry versus electronic trading. Journal of International Financial Markets, Institutions, and Money 8, 243–260.
Massimb, M.N., Phelps, B.D., 1994. Electronic trading, market structure and liquidity. Financial Analysts Journal 50, 39–50.
Osterwald-Lenum, M., 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics. Oxford Bulletin of Economics and Statistics 54, 461–472.
Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in a time series regression. Biometrika 75, 335–346.
Silber, W.L., 1984. Marketmaker behavior in an auction market: an analysis of scalpers in futures markets. Journal of Finance 39 (4), 937–953.
Stock, J., Watson, M.W., 1988. Testing for common trends. Journal of the American Statistical Association 83, 1097–1107.