Yang-Cheng Lu (Taiwan), Yu-Chen Wei (Taiwan)

Classification of trade direction for an equity market with price limit and order match: evidence from the Taiwan stock market

Abstract

This study investigates the applicability and accuracy of revised trade direction algorithms to the Taiwan Stock Exchange (TWSE) data including the tick rule, reverse tick rule, quote rule, at the quote rule, revised quote rule, the Lee and Ready (LR) algorithm, and the Ellis, Michaely and O'Hara (EMO) algorithm. While there are price limits and no designated market maker with the order matching system in the TWSE, we propose that the appropriate classification rule for the TWSE should first adjust the 'no bid or no offer quote' problems, and then classify a trade according to the quote rule and finally the tick rule. We refer to this as a revised LR algorithm (henceforth the RLR algorithm). The empirical results show that there is almost 59% of trades at the zero tick and 94% of trades at the quote, which supports the notion that the quote rule should be identified before the tick rule in the TWSE. The analysis of the other classifications shows that the degree of accuracy ranges from 67.11% to 96.89% compared with the RLR algorithm. The RLR algorithm proposed in this paper could be applied to other topics related to the market microstructure and the empirical results could also be applied to other emerging markets with price limits and order matching systems.

Keywords: buyer/seller initiated trades, microstructure, Lee and Ready algorithm, price limit, order match, Taiwan stock exchange.
JEL Classification: G12, G11, G14.

Introduction

The classification of trades is a major and fundamental subject within the framework of the information content of trades, the order imbalance and inventory accumulation of liquidity providers, the price impact of large trades, the effective spread, and many other related issues. Hasbrouck (1988) showed that the classification of trades as buys or sells is used to test asymmetric-information and inventory-control theories of specialist behavior. Blume, MacKinlay, and Terker (1989) posited that a buy-sell classification is used to measure order imbalance in tests of breakdowns in the linkage between S&P stocks and non-S&P stocks during the crash of October, 1987. In Harris (1989), an increase in the ratio of buys to sells is used to explain the anomalous behavior of closing prices. Lee (1990) showed that the imbalance in buy-sell orders is used to measure the market response to an information event. In Holthausen, Leftwich, and Mayers (1987), a buy-sell classification is used to examine the differential effect of buyer-initiated and seller-initiated block trades. All the previous studies apply the buy-sell classification methods to proceed with the analysis.

Intraday databases of stock exchanges do not provide information on the true buyer/seller initiated trade direction. Consequently, empirical researchers have relied on trade direction algorithms to classify trades as being either buyer or seller motivated. The pioneering work of Lee and Ready (1991) evaluates alternative methods for classifying individual trades as market buy or market sell orders using intraday trade and quote data for a sample of 150 NYSE firms during 1988. They recommended that a combination of quote and tick algorithms be used in practice (hereafter referred to as the LR algorithm).

There are various studies that assess the accuracy of algorithms to infer the direction of trade using the TORQ sample of NYSE trades. The TORQ dataset includes trading information on 144 NYSE stocks for a three-month period beginning in November 1990. Lee and Radhakrishna (2000) use TORQ to calibrate several techniques commonly employed to infer investor behavior from transactions data. They evaluate the LR algorithm to determine the direction of trade, and examine the use of trade size as a proxy for the trader's identity. For those trades that can be classified, the LR algorithm is found to be 93% accurate. They also construct a firm-specific trade size proxy that is highly effective in separating the trading activities of individual and institutional investors. Odders-White (2000) further employs the TORQ data to investigate the performance of the Lee and Ready (1991) trade classification algorithm. Odders-White (2000) finds that the LR algorithm systematically misclassifies transactions at the midpoint of the bid-ask spread, not only small transactions, but also transactions in large or frequently-

This work is supported by the project of the National Science Council in Taiwan (Grant number: NSC 93-2416-H-130-008 and NSC 94-2416-H-130-012).

traded stocks. Finucane (2000) also uses data from the NYSE's TORQ database to test the ability of several competing methods (the tick test, LR (1991) algorithm, and the reverse tick test) to identify market buy and sell orders using intraday quote and trade prices, and identifies factors affecting the accuracy of the methods. These studies all indicate that the LR algorithm biases are systematic. Trades in more liquid stocks, and those involving smaller amounts, tend to be misclassified more frequently.

Three other studies directly examined the accuracy of classification methods using non-NYSE data. Aitken and Frino (1996) test the tick test's accuracy by comparing its predicted direction with the actual direction of trade for a sample of Australian Stock Exchange trades. They also indicate that attempting to apply the tick rule when the best quotes are moving will bias classification in the direction of the market movement. Ellis, Michaely and O'Hara (2000) study the accuracy of the quote, tick and Lee and Ready methods using NASDAQ data that contain 313 stocks traded between September 27, 1996, and September 29, 1997. They also propose a new and simpler classification algorithm, which uses the quote rule to classify trades at the quote (bid or ask) and the tick rule to classify all other trades. Theissen (2001) analyzes the accuracy of the LR (1991) trade classification algorithm and the tick test for a sample taken from the Frankfurt Stock Exchange which is the first paper to use data from a European market. The LR method classifies 72.8% of the transactions correctly. However, the simpler tick test performs almost equally well. He also documents that the misclassification of trades may systematically bias the results of empirical microstructure research.

The validity of many economic studies hinges on the ability to properly classify trades as either buyer- or seller-initiated (Odders-White, 2000). Boehmer, Grammig, and Theissen (2006) use order data from the NYSE and find that inaccurate trade classification algorithms lead to downward bias in estimates of the probability of informed trading. It is, therefore, essential for a reliable classification algorithm to be established. Although there are various kinds of classification rules, different exchanges have suitable rules which can describe the properties of their respective trading system. Most studies concentrate on the NYSE and NASDAQ, an auction market and a dealer market, respectively. Aitken and Frino (1996) focus on the Australian Stock Exchange (ASX), which uses the Stock Exchange Automated Trading System (SEATS). The SEATS is primarily an order-matching system. This contrasts with the London Stock Exchange and the NYSE, which are primarily quote-driven markets in which market-makers /specialists play a prominent role.

The Taiwan Stock Exchange (TWSE) is primarily an order-driven system with price limits and no market makers, which is a similar system to the Australian Stock Exchange (ASX). A market maker is responsible for ensuring that a market is available for listed securities by posting a bid and ask price. On the NASDAQ stock exchange, market maker is required to provide a "two-sided quote" for the securities they cover. Since there is no market maker in the Taiwan stock market, the situation of "no bid or no offer quote" commonly appears in the Taiwan stock market when the liquidity is low or when the price limit is reached. If there is only a bid (ask) price for a security, it might be classified as a sell (buy) trade based on the Lee and Ready algorithm. In this case, the trade is misclassified and it should be classified as a buyer- (seller-) initiated trade since there is only a buy (sell) side quote as a result of the liquidity problem. For this reason, there is a need to investigate an appropriate trade classification rule for the TWSE.

Analyzing the accuracy of the trade classification is of obvious importance because such accuracy determines the validity of empirical research based on the classification algorithm. Analyzing accuracy, though, requires knowledge of the "true trade classification" (Theissen, 2001). Odders-White (2000) studies the TORQ dataset and points out that the initiator of a transaction is the investor (buyer or seller) who has placed his or her order last, chronologically. Theissen (2001) investigates the Frankfurt Stock Exchange and notes the true trade classification as based on whether the Makler (the equivalent of a specialist on the Frankfurt Stock Exchange) has bought or sold shares. If the Makler sold (bought) shares, the transaction is classified as being buyer-initiated (seller-initiated). This is similar to the approach of Ellis, Michaely and O'Hara (2000). They analyze the NASDAQ and classify a trade as being buyer-initiated (seller-initiated) if a customer or broker bought shares from (sold shares to) a market-maker or if a customer bought shares from (sold shares to) a broker. Inter-broker and inter-dealer trades are not classified. The TWSE, by contrast, does not provide data on the buyer-/seller-initiated trade direction on the trade file, order file, or disclosure file. We therefore propose that the appropriate trade classification rule that is applied in the order match system for the TWSE adjusts the identification of the "only bid or only ask price" trade. For this reason, we investigate the appropriate trade direction classification for the TWSE and further compare different classification rules for that exchange.

In this paper, we investigate the applicability and accuracy of revised trade direction algorithms to the Taiwan Stock Exchange (TWSE). To summarize, our analysis focuses on resolving the following issues. First, we investigate an appropriate trade classification algorithm for the TWSE. Second, we summarize the buyer-/seller-initiated trade classification for sub-samples of trades based on differences in price movements and trade sizes. Finally, we analyze the degree of success as a result of the different classification rules for the TWSE by comparing the rules with the appropriate classification algorithm proposed in this study. The remainder of this paper is organized as follows. Section 1 illustrates the methods used to infer the trade direction used in this study. Section 2 describes the data. Section 3 presents the results of the classification. The last section concludes.

1. Methods of inferring trade direction

1.1. Appropriate trade classification for the TWSE. The Taiwan stock market is an order-driven market that differs from an auction markets such as the NYSE or a dealer markets such as NASDAQ. In the Taiwan stock market, there is no market maker and therefore the "only bid or only ask price" is sometimes quoted for particular securities. A "no bid or no ask price" may commonly appear when the market liquidity is low. The misclassification is summarized in Figure 1.

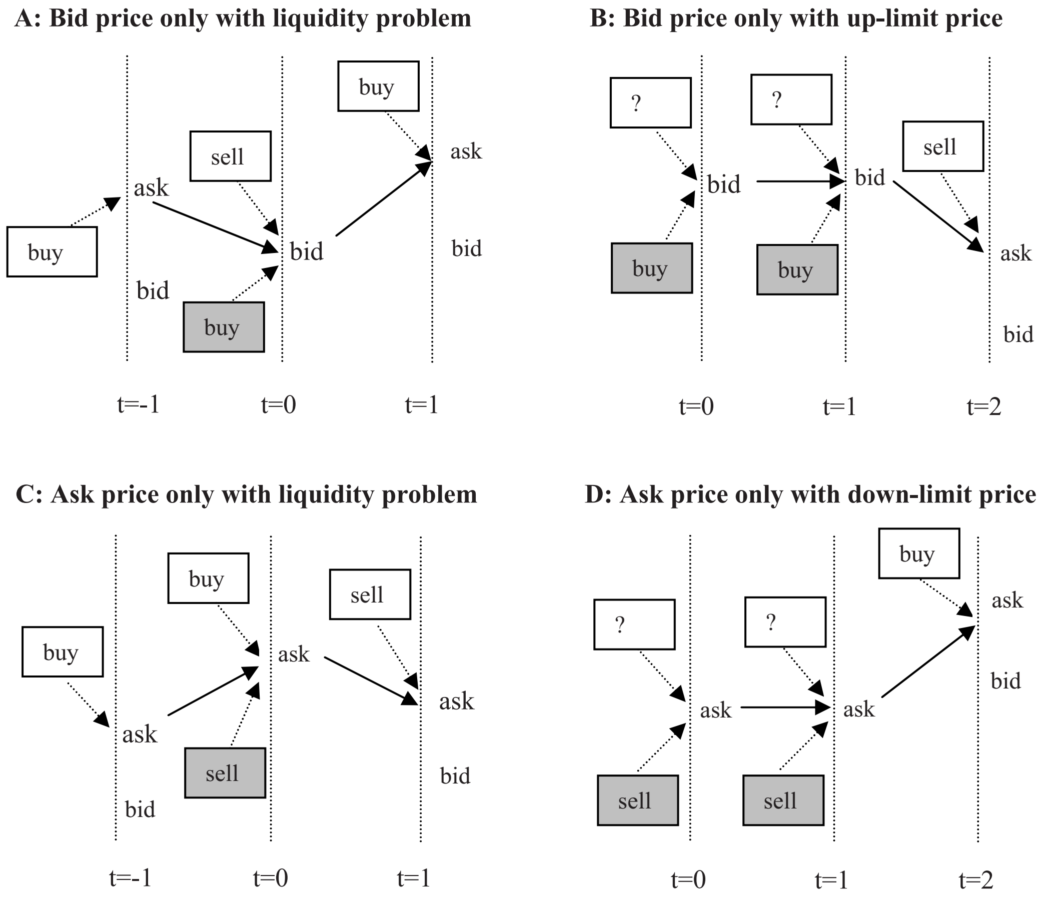

Fig. 1. Misclassification types on the TWSE considering the liquidity and price limit

Note: Panel A presents the case that there is only a bid price and that the trade price does not reach the up-limit price. Panel B presents the case where there is only a bid price and the trade price reaches the up-limit price. Panel C presents the case where there is only an ask price and that the trade price does not reach the down-limit price. Panel D presents the case where there is only an ask price and that the trade price reaches the down-limit price. The solid line conveys the price traded at the bid or ask price at time t. The rectangle with the dotted line refers to the buyer-/seller-initiated trade classification, based on the Lee and Ready algorithm at time t. The question mark indicates that the trade could not be classified based on the Lee and Ready algorithm in the opening trading session. The shading in the rectangle means that there is a misclassification of that trade due to the "no bid or no ask price" problems and that the buy or sell in the shaded rectangles refers to the appropriate buyer-/seller-initiated trade classified at time t. The misclassification in panels A and C is the trade at time t=0 due to the liquidity problem, and the misclassification in panels B and D is the trade at time t=0 and t=1 owing to the price limit.

There are four cases where the application of the trade classification may mis-classify the trade direction after taking the "no bid or no ask price" problem into consideration. First, there is only a bid price in the market and the trade price does not reach the up-limit price, which is indicated in panel A of Figure 1. Since there is no designated market-maker for the TWSE, this case would occur due to the liquidity problem. Second, there is only a bid price in the market because the trade price reaches the up-limit price and is shown in panel B of Figure 1. This case would occur because of the regulation that imposes price limits for the TWSE. Third, there is only an ask price and the trade price is not at its down-limit price which is the case in panel C of Figure 1. The reason why this particular case would occur is the same as in the first case because of the liquidity problem. The fourth case is that where there is only an ask price because the trade price reaches the down-limit price, and it is shown in panel D of Figure 1.

In the first and the second cases, the trade price would be misclassified as a sell while using the tick, reverse tick, and LR algorithm. On the other hand, these cases would not be classified using the quote rule since there is no ask price. These cases should, therefore, be classified as being buyer-initiated since the trade is complete owing to the bid buy side orders. The third and fourth cases are misclassified for the same reasons as in the first two cases. The third and fourth cases should, therefore, be classified as being seller-initiated since the trade is complete owing to the offer sell side orders.

The TWSE is primarily a pure order-matching system in which there is no designated market maker; therefore, the true-trade classification rules used in Odders-White (2000), Theissen (2001) and EMO (2000) are not appropriate for the TWSE. Our study finds that the quote rule can classify more than 90% of the price changes in the TWSE, which will be shown in empirical analysis later. On the other hand, the appropriate trade classification in the TWSE should reflect the 'no bid or no ask price' problem, which might be caused by the liquidity and price limit. We therefore use the revised quote rule, which adjusts the "no bid or no ask price" problem before the quote rule to classify trades, and then the tick rule to classify all other trades. The LR algorithm adopts the quote rule followed by the tick rule; therefore, the trade-direction classification rule proposed in our study could be considered to be a revised LR method (henceforth the RLR algorithm). We propose that the RLR algorithm is an appropriate classification rule for the TWSE. We then compare different trade classification rules for the TWSE to further confirm that the RLR algorithm could classify almost 100% of the trades on the TWSE, an order-match system with price limits and no designated market maker.

1.2. Competing methods to identify the trade direction. Our study tests the ability of several competing methods to identify market buy and sell orders. The most commonly-used methods to infer the trade direction are the tick rule, reverse tick, and LR (1991) algorithm. In considering the recent related studies, we apply the rules used in EMO (2000). The content of each classification rule is described as follows:

The tick rule. The tick rule is based on price movements relative to previous trades. If the transaction is above (below) the previous price, then it is a buy (sell). If there is no price change but the previous tick change was up (down), then the trade is classified as a buy (sell).

The reverse tick. The reverse tick test uses the next trade price to classify the current trade. If the next trade occurs on an uptick or zero uptick, the current trade is classified as a sell. If the next trade occurs on a downtick or zero downtick, the current trade is classified as a buy.

The quote rule. The quote rule classifies a transaction as a buy if the associated trade price is above the midpoint of the bid and ask; it is classified as a sell if the trade price is below the midpoint quote.

The at the quote rule. The at the quote rule classifies a transaction as a buy if the associated trade price is traded at the asking price; it is classified as a sell if the trade price is at the bidding price.

The revised quote rule. In considering the problems of a "no bid or no offer quote" in the TWSE, we proposed a revised quote rule that considers the adjustment of the price limit before the quote rule. The trade would be classified as a buy if there is only the bid-side quote and it would be classified as a sell if there is offer-side quote only.

The LR algorithm. The LR algorithm (Lee and Ready, 1991) is essentially a combination of these two rules: first, classify a trade according to the quote rule (above or below the midpoint), and then classify the midpoint transaction using the tick rule. In considering the reporting procedure on the NYSE, Lee and Ready also suggest comparing transaction prices with quotes reported at least five seconds before the transaction is reported. Since the adjustment of the "five seconds" before the transaction could not be implemented on the TWSE, we would disregard the comparison of the "five seconds" before the transaction and apply just the current price of the quote and tick.

The EMO algorithm. Ellis, Michaely, and O'Hara (2000) (simplified as the EMO (2000) algorithm) use the quote rule to classify trades at the quote (bid or ask) and the tick rule to classify all other trades, which means that the EMO method classifies the trades by means of the "at the quote" rule first and then the tick rule.

The tick rule and the reverse tick rule could deal with almost every possible trade. There is, however, misclassified or incorrect trade identification in the case of no bid or no ask price in the (opening) trading session. According to the quote rule and at the quote rule, the trade could be classified only if both the bid and ask prices exist; therefore, the quote rule and at the quote rule might be misclassified if there is only a bid or ask price. The classification algorithm that combines the tick or quote rule with other rules might experience such a misclassification.

2. The TWSE database

The TWSE is primarily an order-matching system similar to the Australian Stock Exchange (ASX). This is contrast with the London Stock Exchange and the NYSE which are primarily quote-driven markets in which market-markers/specialists play a prominent role. We introduce the trading of securities in the TWSE and the database we used in the study in the following sections.

2.1. Trading of securities on the TWSE. When the TWSE was first established, trading in the centralized market was carried out in an open-outcry manner. In order to keep abreast of the changing needs of the market environment, the trading procedure has progressed through several evolutionary phases. In August 1985, the open-outcry system was gradually replaced by a computer-aided trading system (CATS), which was eventually upgraded to a fully automated securities trading (FAST) system in 1993. The centralized market trading session lasts from 9:00 A.M. to 1:30 P.M., Monday through Friday (with some Saturdays adjusted to trade being included)[1]. (Orders can be entered from 8:30 A.M. to 1:30 P.M.) The off-hour trading session is 2:00-2:30 P.M., Monday through Friday. Investors may place an order in person, by phone, fax or through the Internet. Orders are entered via terminals on securities firms' premises into the TWSE's main computer and are processed and executed by the trading system on a price-and-time-priority principle. In special cases, listed stocks may be traded through negotiation, auction, tender, or other means.

There are two types of matching method by continuous auction, which are as follows:

(1) For a single security: The time/price priority for matching shall be based on the following principles whenever a buying or selling order enters the system: Incoming buying (selling) orders whose prices are greater (less) than or equal to the lowest (highest) previously entered buying (selling) orders will be matched and executed at the individual ask (bid) prices sequentially from the lowest (highest) to the highest (lowest); if two or more quotes show identical bid (ask) prices, they will be matched and executed sequentially in chronological order until all buying (selling) orders are satisfied or until the price of the current incoming buying (selling) order is lower (higher) than the bid (ask) prices of unexecuted selling (buying) orders.

(2) For a basket of stocks: The stock codes, unit prices, and volumes of incoming buying (or selling) orders shall all be identical to those of the previously entered selling (or buying) orders; the orders are then executed with selling (or buying) orders sequentially in chronological order.

Trading prices are decided by call auction. The TWSE conducts "intra-day volatility interruption" to prevent the over-volatility of stock prices, and also discloses the prices and volume of unexecuted orders for the five best bids/asks. At the end of the trading session, the trading system accumulates orders for five minutes (from 1:25 P.M. to 1:30 P.M.) before the closing call auction, in order to form fair closing prices.

Like the other emerging markets such as Korea and China, some price limit regulations are set on the TWSE. There are daily price limits and the minimum up/down tick size of price movements for the stocks traded on the TWSE, excluding the first five trading days after a listing. The price limit of the stock is the positive and negative 7% of the previous day's closing price, which is different from the exchanges in developing markets, such as the NYSE and NASDAQ. Table 1 presents the annual statistics for the TWSE. The listed companies grow rapidly during the period from 1997 through 2006. The trading percentage of foreign investors also increased to 18%~19% in the past two years.

2.2. The TWSE's database of empirical studies. The sample contains 684 TWSE stocks traded from January 2, 2006 through June 30, 2006, excluding mutual funds, warrants, and corporate bonds. The transactions data are provided by the Taiwan Stock Exchange (TWSE). Overall, the sample is taken from 120 trading days and 17,272,235 trades.

[1] The regular trading session lasts from 9:00 A.M. to 1:30 P.M., Monday through Friday. Saturdays may be adjusted for trading if there are holidays on the regular trading days. The Central Personnel Administration in Taiwan will announce adjustments to trading on Saturdays if it is necessary.

Table 1. Description of data for the Taiwan Stock Exchange

Year	No. of listed companies	Trading days	Total market value at year-end (NT $Million)	Trading percentages of investor types (%)			
				Domestic individuals	Domestic institutions	Foreign individuals	Foreign institutions
1997	404	286	9,696,113	90.73	7.55	0.01	1.71
1998	437	271	8,392,607	89.73	8.63	0.02	1.62
1999	462	266	11,803,524	88.23	9.36	0.01	2.4
2000	531	271	8,191,474	86.1	10.27	0.01	3.62
2001	584	244	10,247,599	84.41	9.69	0.01	5.89
2002	638	248	9,094,936	82.3	10.05	0.97	6.68
2003	669	249	12,869,101	77.84	11.51	1.24	9.41
2004	697	250	13,989,100	75.94	11.56	1.63	10.87
2005	691	247	15,633,858	68.84	13.29	2.41	15.46
2006	688	248	19,376,975	70.56	11.04	2.25	16.15

Note: The data source is the Taiwan Stock Exchange.

The transactions data are preserved in the following three files: the order file, trade file, and disclosure file. The trade file includes the date, stock code, trade time, order type (buy or sell) of transaction, trade volume, trade serial number, trade price, trade categories[1], and the identity of the trader. The order file contains the date, stock code, order type (buy or sell) of the transaction, trade categories, trade time, identity of the trader, and so on. The disclosure file illustrates the trade price, disclosure price of the bid, ask and the date. The identity of the trader includes mutual funds, foreign investors, individual investors, dealers, and general institutional investors. The intraday files provide the details of the trade, order, and disclosure information. The true buyer/seller-initiated trade direction is not disclosed in the intraday information. That is why we propose the appropriate trade classification algorithm on the Taiwan Stock Exchange with price limits and order-match system.

3. Classification results

Table 2 presents the summary statistics of daily price movements, including the price traded at the midpoint, the inside spread, at the quotes, no bid or no offer quote, and other outside the quotes. The "no bid or no offer quote" averages almost 1.6%. Although the percentage seems low, this problem can not be ignored in the TWSE with price limits and no market maker. On the other hand, we can see that prices traded at the midpoint are at most 3.12%, which means that the remaining trades can be classified by the quote rule with the adjustment of the "no bid or no offer quote". That is why we apply the "revised quote rule" to classify the trade and then the tick rule as the appropriate classification rule in the TWSE.

Table 2. Summary statistics of daily trade location percentages

	Midpoint	Inside spread	At the quotes	Bid price or ask price only	Outside the quotes
Mean (%)	3.1176	2.5156	92.7098	1.6179	0.0390
Median (%)	3.1250	2.3800	92.8650	1.4050	0.0400
Maximum (%)	4.2100	4.5000	94.5300	5.6500	0.0900
Minimum (%)	2.3400	1.7000	87.7900	0.2400	0.0200
Std. dev.	0.3701	0.5752	1.0997	1.0073	0.0149
Skewness	0.0895	1.1132	-1.1472	1.2270	1.0670
Kurtosis	2.5241	3.9575	5.4606	4.8183	3.9306
Jarque-Bera	1.2929	29.3677	56.5931	46.6412	27.0996

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. The trade location includes a trade at the midpoint, the inside spread where the trade is between the bid and ask but not at the midpoint, a trade at the quotes, where the trade occurs when there is only a bid side quote or an offer side quote, and a trade in the other outside the quotes situation.

[1] The trade categories include spot transaction, margin long, and margin short.

Table 3 presents the classificatory power of different trade rules for tick direction including the downtick, zero tick, and uptick. A total of 59.24% of the trades occur on zero ticks, 20.74% on downticks, and 20.01% on upticks. The RLR algorithm classifies almost all of the zero tick trades, and the other rules classify from 55.24% to 57.98% of the trades on the zero ticks. We also find that the seller-initiated percentage is larger than the buyer-initiated percentage in the trade classification rules except for the reverse tick test. In general, the RLR algorithm is able to classify nearly 100% of the trades during our study period, while the other rules are only able to classify from 93.13% to 98.74% of the trades. These empirical results support the expectation of our study that the revised LR algorithm would be a more appropriate classification rule for the TWSE.

Table 3. Summary of buyer/seller-initiated trades for competitive classification rules based on the tick and trade direction

Classification	Full sample			Tick direction								
				Downtick			Zero tick			Uptick		
	B	S	T	B	S	T	B	S	T	B	S	T
Tick rule	47.42	51.32	98.74	0.00	20.74	20.74	27.41	30.57	57.98	20.01	0.00	20.01
Reverse tick rule	50.50	47.51	98.02	7.10	13.40	20.50	29.97	27.75	57.73	13.42	6.36	19.79
Quote rule	46.03	47.10	93.13	3.06	15.93	18.99	27.28	27.96	55.24	15.69	3.22	18.91
At the quote rule	43.50	50.91	94.41	2.91	16.58	19.49	25.42	30.86	56.28	15.17	3.47	18.64
Revised quote rule	45.63	51.25	96.89	3.05	17.06	20.11	26.89	30.53	57.43	15.69	3.66	19.34
LR algorithm	45.33	52.85	98.18	3.06	17.68	20.74	25.93	31.50	57.43	16.35	3.67	20.01
EMO algorithm	44.87	52.17	97.04	2.91	17.84	20.74	25.42	30.86	56.28	16.54	3.47	20.01
RLR algorithm	47.30	52.70	100.00	3.05	17.69	20.74	27.89	31.35	59.24	16.36	3.66	20.01
Total	17,272,235			3,582,767			10,232,690			3,456,778		
	100.00			20.74			59.24			20.01		

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. This table presents the percentages of buyer/seller-initiated trades for space consideration. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The trade numbers in each category are available from the authors upon request. The trade direction of B (S) represents the buyer- (seller-) initiated trade. T represents the subtotal of each category.

Table 4 provides results for sub-samples of trades based on whether the midpoint of the spread increased or decreased from the open to the close of trade (upward, downward and zero movements respectively). The percentage for no quote changes is approximately 66.41%, which confirms that the percentage of zero movement (67.8%) is larger than the upward or downward movements (15.78% and 16.42% respectively). Similar to the results of Table 3, the RLR algorithm classifies almost 100% of the trades, whereas other rules classify 94.41% to 98.78% of the trades.

Table 4. The percentages of buyer/seller-initiated trades for sub-samples of trades based on a day's price movements and quote change

Classification	Quote change or not				Midpoint upward or not						Total
	Quote change		No quote change		Upward		Downward		Zero		
	B	S	B	S	B	S	B	S	B	S	
Tick rule	16.46	16.93	30.96	34.39	12.87	2.81	2.94	13.38	31.62	35.13	98.74
Reverse tick rule	17.27	15.81	33.23	31.70	7.10	8.45	9.47	6.70	33.93	32.37	98.02
Quote rule	14.91	14.67	31.12	32.44	5.05	8.57	9.33	5.49	31.65	33.05	93.13
At the quote rule	13.82	14.86	29.68	36.05	4.44	9.04	8.96	5.17	30.11	36.70	94.41
Revised quote rule	14.87	15.98	30.77	35.27	5.05	9.42	9.30	5.82	31.29	36.01	96.89
LR algorithm	15.41	16.55	29.92	36.30	5.60	9.44	9.28	6.33	30.45	37.09	98.18
EMO algorithm	15.01	15.97	29.87	36.20	5.53	9.07	8.99	6.14	30.36	36.95	97.04
RLR algorithm	16.33	17.26	30.97	35.44	5.96	9.82	9.79	6.62	31.55	36.26	100.00

Table 4 (cont.). The percentages of buyer/seller-initiated trades for sub-samples of trades based on a day's price movements and quote change

Total	5,802,291	11,469,944	2,725,224	2,835,645	11,711,366	
	33.59	66.41	15.78	16.42	67.80	

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. This table presents the percentages of buyer/seller-initiated trades for space consideration. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The trade numbers in each category are available from the authors upon request. The trade direction of B (S) represents the buyer- (seller-) initiated trade.

Table 5 presents the summary for sub-samples of the classification of trades based on different price changes, including trades at the midpoint, inside spread, at the quote (bid or ask), and the outside the quotes condition. The results show that 94.41% of the trades are transacted at the quote, 3.07% at the midpoint, 2.48% at the inside spread, and 0.04% at the outside the quotes. The total zero tick is approximately 59.24% in Table 3 and the trade at the quotes is almost 94.41% in Table 5; therefore, it confirms that the quote rule should be applied before the tick rule no matter what the algorithms are applied for the TWSE.

Table 5. Summary of buyer/seller-initiated trades for competitive classification rules based on different price changes

Classification	Midpoint		Inside spread		At the quotes		Outside the quotes		Total
	B	S	B	S	B	S	B	S	
Tick rule	1.52	1.52	1.23	1.19	44.66	48.60	0.02	0.01	98.74
Reverse tick rule	1.70	1.35	1.34	1.11	47.45	45.03	0.02	0.02	98.02
Quote rule	0.00	0.00	1.17	1.31	44.85	45.79	0.01	0.00	93.13
At the quote rule	0.00	0.00	0.00	0.00	43.50	50.91	0.00	0.00	94.41
Revised quote rule	0.00	0.00	1.16	1.31	44.47	49.94	0.00	0.00	96.89
LR algorithm	0.67	0.63	1.16	1.31	43.50	50.91	0.00	0.00	98.18
EMO algorithm	0.67	0.63	0.70	0.63	43.50	50.91	0.00	0.00	97.04
RLR algorithm	1.64	1.42	1.17	1.32	44.47	49.94	0.02	0.02	100.00
Total	529,634		428,981		16,306,858		6,762		17,272,235
	3.07		2.48		94.41		0.04		100.00

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. This table presents the percentages of buyer/seller-initiated trades for space consideration. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The trade numbers in each category are available from the authors upon request. The trade direction of B (S) represents the buyer- (seller-) initiated trade.

Does the size of a trade affect the likelihood of correctly classifying it as a buy or sell? Table 6 shows the distribution of the percentages of buyer-/seller-initiated trades based on trade size. The results suggest that the seller-initiated percentage is larger than the buyer-initiated percentage of trade classification rules in most of the trade size decile except for the reverse tick test which confirms the finding in Table 3. Overall, there is a monotonic relationship: a better classification for smaller trades, which indicates that there are larger trades in the smaller trade size in the TWSE.

Table 6. Distribution of the percentages of buyer/seller-initiated trades by trade size

Classification		Trade size decile										Total
		Small	20%	30%	40%	50%	60%	70%	80%	90%	Large	
Tick rule	B	39.69	4.46	1.50	0.64	0.31	0.28	0.13	0.08	0.06	0.28	47.42
	S	42.40	5.10	1.72	0.74	0.36	0.33	0.16	0.10	0.07	0.35	51.32
Reverse tick rule	B	42.21	4.88	1.62	0.68	0.33	0.29	0.13	0.08	0.06	0.22	50.50
	S	39.26	4.70	1.62	0.71	0.35	0.31	0.15	0.09	0.06	0.26	47.51
Quote rule	B	38.31	4.39	1.49	0.65	0.32	0.28	0.14	0.09	0.06	0.30	46.03
	S	38.96	4.61	1.56	0.68	0.33	0.31	0.15	0.09	0.07	0.35	47.10

Table 6 (cont.). Distribution of the percentages of buyer/seller-initiated trades by trade size

Classification		Trade size decile										Total
		Small	20%	30%	40%	50%	60%	70%	80%	90%	Large	
At the quote rule	B	36.33	4.12	1.39	0.60	0.30	0.25	0.12	0.08	0.06	0.26	43.50
	S	42.39	4.84	1.63	0.71	0.35	0.31	0.15	0.10	0.07	0.36	50.91
Revised quote rule	B	37.98	4.36	1.48	0.65	0.32	0.28	0.13	0.08	0.06	0.29	45.63
	S	42.37	5.01	1.71	0.74	0.36	0.34	0.16	0.10	0.07	0.38	51.25
LR algorithm	B	37.62	4.39	1.49	0.65	0.32	0.28	0.13	0.08	0.06	0.29	45.33
	S	43.71	5.14	1.76	0.76	0.38	0.35	0.17	0.11	0.08	0.40	52.85
EMO algorithm	B	37.31	4.32	1.47	0.64	0.31	0.27	0.13	0.08	0.06	0.29	44.87
	S	43.24	5.03	1.71	0.74	0.37	0.34	0.16	0.10	0.07	0.39	52.17
RLR algorithm	B	39.29	4.55	1.55	0.67	0.33	0.29	0.14	0.09	0.06	0.31	47.30
	S	43.50	5.19	1.77	0.77	0.38	0.35	0.17	0.11	0.08	0.40	52.70
Total		14,298,476	1,683,021	574,092	249,356	122,571	111,821	53,194	33,574	24,454	121,676	17,272,235
		82.78	9.74	3.32	1.44	0.71	0.65	0.31	0.19	0.14	0.70	100.00

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. This table presents the percentages of buyer/seller-initiated trades for space consideration. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The trade numbers in each category are available from the authors upon request. The trade direction of B (S) represents the buyer- (seller-) initiated trade.

Table 7 provides a summary of trade classifications compared with the trade size deciles and the price changes including the midpoint, inside spread, at the quotes, and outside the quotes. No matter what the trade rule is applied, price movements of the at the quotes deliver the larger buyer/seller-initiated classification in each trade size decile from the small to the large. The results also confirm that most of the trades are classified at the quotes (trade at bid or ask) and the smaller trade decile. To sum up, the empirical results of Table 6 and Table 7 support the notion that most of the trades are in the small trade size decile in TWSE. Besides, the findings that most of trades are classified at the quotes further provide the robustness check that the quote rule should be applied before the tick rule in TWSE.

Finally, Table 8 compares the performance of different trade classifications with the appropriate RLR algorithm for the TWSE. The reverse tick rule identifies 67.11% of the trades, which is the lowest rate of accuracy; and the tick rule achieves a 74.18% rate of accuracy, being the second lowest rate of accuracy. Classification rules which consider the quote change first provide higher accuracy such as the quote rule, the at the quotes rule, the revised quote rule, the LR algorithm, and the EMO algorithm. The rules also show that the rates of unclassified and misclassified data will be higher if the trade or quote rule is taken into account alone. The results indicate that the LR method performs slightly better than the EMO approach, since the LR algorithm applies the quote rule first, while the EMO's algorithm uses the "at the quote rule" before the tick rule. The difference between the quote rule and the "at the quote rule" lies in the quote rule is capable of classifying the inside-spread trades.

Table 7. Summary of trade classifications compared with trade size decile and price movements

Classification			Trade size decile										Total
			Small	20%	30%	40%	50%	60%	70%	80%	90%	Large	
Tick rule	midpoint	B	1.21	0.18	0.06	0.03	0.01	0.01	0.01	0.00	0.00	0.01	1.52
		S	1.17	0.19	0.07	0.03	0.01	0.02	0.01	0.00	0.00	0.01	1.52
	inside	B	0.83	0.20	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.23
		S	0.77	0.20	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.19
	at the quotes	B	37.64	4.08	1.36	0.58	0.28	0.24	0.12	0.07	0.05	0.25	44.66
		S	40.45	4.71	1.56	0.67	0.32	0.29	0.14	0.09	0.06	0.31	48.60
	outside	B	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02

Table 7 (cont.). Summary of trade classifications compared with trade size decile and price movements

Classification			Trade size decile										Total
			Small	20%	30%	40%	50%	60%	70%	80%	90%	Large	
		S	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Reverse tick rule	midpoint	B	1.33	0.21	0.07	0.03	0.01	0.02	0.01	0.00	0.00	0.01	1.70
		S	1.06	0.16	0.06	0.03	0.01	0.01	0.01	0.00	0.00	0.01	1.35
	inside	B	0.89	0.22	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.34
		S	0.73	0.18	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.11
	at the quotes	B	39.98	4.45	1.45	0.61	0.29	0.25	0.12	0.07	0.05	0.18	47.45
		S	37.46	4.36	1.49	0.65	0.32	0.27	0.13	0.08	0.06	0.22	45.03
	outside	B	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
		S	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Quote rule	midpoint	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	inside	B	0.78	0.19	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.17
		S	0.85	0.22	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.31
	at the quotes	B	37.52	4.19	1.42	0.61	0.31	0.25	0.13	0.08	0.06	0.28	44.85
		S	38.11	4.39	1.47	0.63	0.31	0.28	0.14	0.09	0.06	0.32	45.79
	outside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
At the quotes rule	midpoint	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	inside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	at the quotes	B	36.33	4.12	1.39	0.60	0.30	0.25	0.12	0.08	0.06	0.26	43.50
		S	42.39	4.84	1.63	0.71	0.35	0.31	0.15	0.10	0.07	0.36	50.91
	outside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Revised quote rule	midpoint	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Inside	B	0.77	0.19	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.16
		S	0.85	0.22	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.31
	at the quotes	B	37.20	4.16	1.40	0.61	0.30	0.25	0.13	0.08	0.06	0.27	44.47
		S	41.52	4.80	1.61	0.70	0.34	0.31	0.15	0.09	0.07	0.35	49.94
	outside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LR algorithm	midpoint	B	0.51	0.08	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.67
		S	0.47	0.08	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.63
	Inside	B	0.77	0.19	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.16
		S	0.85	0.22	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.31
	at the quotes	B	36.33	4.12	1.39	0.60	0.30	0.25	0.12	0.08	0.06	0.26	43.50
		S	42.39	4.84	1.63	0.71	0.35	0.31	0.15	0.10	0.07	0.36	50.91
	outside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 7 (cont.). Summary of trade classifications compared with trade size decile and price movements

Classification			Trade size decile										Total
			Small	20%	30%	40%	50%	60%	70%	80%	90%	Large	
EMO algorithm	midpoint	B	0.51	0.08	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.67
		S	0.47	0.08	0.03	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.63
	inside	B	0.46	0.12	0.05	0.02	0.01	0.02	0.01	0.00	0.00	0.02	0.70
		S	0.38	0.11	0.05	0.02	0.01	0.02	0.01	0.00	0.00	0.02	0.63
	at the quotes	B	36.33	4.12	1.39	0.60	0.30	0.25	0.12	0.08	0.06	0.26	43.50
		S	42.39	4.84	1.63	0.71	0.35	0.31	0.15	0.10	0.07	0.36	50.91
	outside	B	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
RLR algorithm	midpoint	B	1.29	0.20	0.07	0.03	0.01	0.02	0.01	0.00	0.00	0.01	1.64
		S	1.10	0.17	0.06	0.03	0.01	0.02	0.01	0.00	0.00	0.01	1.42
	inside	B	0.78	0.19	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	1.17
		S	0.85	0.22	0.09	0.04	0.02	0.03	0.01	0.01	0.00	0.03	1.32
	at the quotes	B	37.20	4.16	1.40	0.61	0.30	0.25	0.13	0.08	0.06	0.27	44.47
		S	41.52	4.80	1.61	0.70	0.34	0.31	0.15	0.09	0.07	0.35	49.94
	outside	B	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
		S	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Total			14298476	1683021	574092	249356	122571	111821	53194	33574	24454	121676	17272235
			82.78	9.74	3.32	1.44	0.71	0.65	0.31	0.19	0.14	0.70	100.00

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. This table presents the percentages of buyer/seller-initiated trades for space consideration. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The trade numbers in each category are available from the authors upon request. The trade direction of B (S) represents the buyer- (seller-) initiated trade.

Table 8. Rate of accuracy of trade rules compared with the RLR algorithm

Classification		True buy		True sell		Rate of accuracy (%)
		Number	Percent (%)	Number	Percent (%)	
Tick rule	buy	5,997,550	34.73	2,193,225	12.70	74.18
	sell	2,049,214	11.86	6,814,177	39.45	
	unclassified	122,306	0.71	95,056	0.55	
Reverse tick rule	buy	5,671,152	32.84	3,051,582	17.67	67.11
	sell	2,286,110	13.24	5,920,591	34.28	
	unclassified	211,808	1.23	130,285	0.75	
Quote rule	buy	7,883,155	45.64	67,065	0.39	92.75
	sell	0	0.00	8,135,936	47.11	
	unclassified	285,915	1.66	899,457	5.21	
At the quotes rule	buy	7,447,790	43.12	65,571	0.38	92.68
	sell	233,250	1.35	8,560,247	49.56	
	unclassified	488,030	2.83	476,640	2.76	
Revised quote rule	buy	7,881,813	45.63	0	0.00	96.89
	sell	0	0.00	8,852,433	51.25	
	unclassified	287,257	1.66	250,025	1.45	

Table 8 (cont.). Rate of accuracy of trade rules compared with the RLR algorithm

Classification		True buy		True sell		Rate of accuracy (%)
		Number	Percent (%)	Number	Percent (%)	
LR algorithm	buy	7,764,400	44.95	65,571	0.38	96.46
	sell	233,250	1.35	8,895,476	51.50	
	unclassified	171,420	0.99	141,411	0.82	
EMO algorithm	buy	7,651,342	44.30	99,282	0.57	94.97
	sell	259,428	1.50	8,750,859	50.67	
	unclassified	258,300	1.50	252,317	1.46	

Note: The full sample consists of 17,272,235 observations during the January 2, 2006, through June 30, 2006, TWSE sample period. The ratio is calculated by the subtotal buyer/seller-initiated trades in each category over the number of total trades during the study period. The rate of accuracy equals the percentage of true buyer-initiated trades adds to the percentage of true seller-initiated trades.

To sum, Table 8 compares and summarizes the rate of accuracy of different classification rules with the RLR algorithm. Since the RLR algorithm adjusts the "no bid or no ask price" problem, it is able to classify almost 100% of the trades. Although the revised quote rule, LR, and EMO algorithms are found to have high rates of accuracy of 96.89%, 96.46% and 94.97%, respectively, there are still classification biases if the "no bid or no ask price" is not adjusted before the quote or tick rule.

Conclusion

The Taiwan Stock Exchange (TWSE) is a pure order-driven market with price limits and no designated market maker so that it differs from the NYSE and the NASDAQ. Since there are price limits and no market marker, the "no bid or no ask price" problem sometimes arises when the securities are quoted. This motivates us to construct an appropriate trade classification rule for the TWSE and to further compare the applicability of the trade direction algorithms to the TWSE data, as well as their accuracy, by considering the tick rule, reverse tick rule, quote rule, at the quote rule, revised quote rule, the LR algorithm, and the EMO algorithm.

Analyzing accuracy of the algorithms requires knowledge of the "true trade classification". Since the TWSE does not declare the true direction of each trade, the definition of true trade classification proposed in previous studies can not be directly applied to Taiwan. Logically, if the "no bid or no ask price" problem can be adjusted before the quote rule and the tick rule, most of the trades on the TWSE would be appropriately identified. For these reasons, we first construct the appropriate RLR trade classification algorithm for the TWSE by adjusting the "no bid or no ask price" problem before the quote and the tick rules been applied. We then compare the different classification rules in a situation where the price varies. While a "no bid or no ask price" may frequently appear in the TWSE, we propose that the "no bid or no ask price" problem should be addressed by focusing on the identification.

The empirical results show that nearly 59.24% of the trades take place at the zero tick and 94.41% of the trades at the quotes. This lends support to the view that the quote rule should be applied before the tick rule in the TWSE, which is the same as in the case of the LR algorithm. The results present that if the "no bid or no ask price" problem can be adjusted before the quote rule and the tick rule applied hereafter, almost 100% of the trades on the TWSE could be identified. The empirical results also confirm that the RLR is applicable to the TWSE. The performances of other algorithms are compared with that of the RLR, and the results show that the reverse tick rule has the lowest rate of accuracy, namely 67.11%, while the revised quote rule has the highest rate of accuracy (96.89%) because the "no bid or no ask price" problem has been adjusted in the revised quote rule. Although previous studies apply the LR algorithm to identify the trade direction in Taiwan, we propose that the RLR algorithm, which makes adjustments for the price limit and the liquidity problem of "no bid or no ask price" could further reflect the realities of the TWSE.

To conclude, the RLR algorithm proposed in this paper could be applied in related studies of market microstructure in emerging markets such as Taiwan, in order to classify trades as buys or sells in the estimation of the probability of information-based trades (PIN). On the other hand, the RLR algorithm could be applied to the data for other emerging markets, especially order-driven markets or markets that have no designated market makers, such as Korea, the Southeast Asian countries, and China.

References

1. Aitken, M., and Frino, A., (1996), "The Accuracy of the Tick Test: Evidence from the Australian Stock Exchange", *Journal of Banking & Finance*, Vol. 20, pp.1715-1729.
2. 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*, Vol. 44, pp. 827-848.
3. Boehmer, E., Grammig, J., and Theissen, E., (2007), "Estimating the Probability of Informed Trading- Does Trade Misclassification Matter?" *Journal of Financial Markets*, Vol. 10, pp. 26-47.
4. Ellis, K., Michaely, R., and O'Hara, M., (2000), "The Accuracy of Trade Classification Rules: Evidence from NASDAQ", *Journal of Financial and Quantitative Analysis*, Vol. 35, pp. 529-551.
5. Finucane, T.J., (2000), "A Direct Test for Methods for Inferring Trade Direction from Intra-Day Data", *Journal of Financial and Quantitative Analysis*, Vol. 35, pp. 553-576.
6. Harris, L., (1989), "A Day-end Transaction Price Anomaly", *Journal of Financial and Quantitative Analysis*, Vol. 24, pp. 29-45.
7. Hasbrouck, J., (1988), "Trades, Quotes, Inventories, and Information", *Journal of Financial Economics*, Vol. 22, pp. 229-252.
8. Holthausen, R.W., Leftwich, R.W., and Mayers, D., (1987), "The Effect of Large Block Transactions on Security Prices: A Cross-sectional Analysis", *Journal of Financial Economics*, Vol. 19, pp. 237-267.
9. Lee, C.M.C., (1990), "Information Dissemination and the Small Trader: An Intraday Analysis of the Small Trader Response to Announcements of Corporate Earnings and Changes in Dividend Policy", Ph.D. dissertation, Cornell University.
10. Lee, C.M.C., and Radhakrishna, B., (2000), "Inferring Investor Behavior: Evidence from TORQ data", *Journal of Financial Markets*, Vol. 3, pp. 83-111.
11. Lee, C.M.C., and Ready, M.J., (1991), "Inferring Trade Direction from Intraday Data", *Journal of Finance*, Vol. 46, pp. 733-746.
12. Odders-White, E., (2000), "On the Occurrence and Consequences of Inaccurate Trade Classification", *Journal of Financial Markets*, Vol. 3, pp. 259-286.
13. Theissen, E., (2001), "A Test of the Accuracy of the Lee/Ready Trade Classification Algorithm", *Journal of International Financial Markets, Institutions and Money*, Vol. 11, pp. 147-165.