The Accuracy of the Tick Rule in the Bitcoin Market

Donglian Ma¹ and Pengxiang Zhai²

Abstract
The tick rule is one of the most popular trade classification algorithms used when an order initiator in market data is not signed. Using 11.9 million trades of Bitcoin/USD on Bitstamp, this article tests the accuracy of the tick rule in the Bitcoin market. Evidence indicates that the overall success rate of the tick rule is 76.87%. It is also shown that the tick rule is inclined to fail in discerning trade intentions when there is a long period of time between trades. Furthermore, order imbalances computed using the tick rule lack sufficient accuracy in the Bitcoin market.

Keywords
Bitcoin, tick rule, classification, accuracy, order imbalance

Introduction
Bitcoin has become the most famous decentralized crypto blockchain since it was first outlined in an article by Nakamoto (2008). Although Bitcoin is originally designed to be a peer-to-peer payment system (Nakamoto, 2008), it has been regarded as a financial asset or an investment tool for a long time. In addition to articles discussing the nature of Bitcoin (Baek & Elbeck, 2015; Baur et al., 2018; Kwon, 2020; Yermack, 2015), a large stream of the literature has explored its financial features, including its price (Baig et al., 2019; Corbet et al., 2018; Hu et al., 2019; Li et al., 2020), return and volatility (Baur et al., 2018; Bouri, Azzi, & Dyhrberg, 2017; Bouri, Jalkh, et al., 2017; Bouri, Molnár, et al., 2017; Dyhrberg, 2016a, 2016b; Klein et al., 2018; Thies & Molnár, 2018), market efficiency (Bariviera, 2017; Kurihara & Fukushima, 2017; Nadarajah & Chu, 2017; Urquhart, 2016), connection with traditional financial markets or within cryptocurrency markets (Andrada-Félix et al., 2020; Antonakakis et al., 2019; Borri, 2019; Kurka, 2019; Li et al., 2020; Nguyen et al., 2020), linkage between price or return and transaction activity on blockchain (Apte & Fiedler, 2020; Koutmos, 2018a, 2018b), and intraday dynamics based on high-frequency data (Bariviera et al., 2018; Eross et al., 2019; Zargar & Kumar, 2019). Most of these studies focus on the performance of Bitcoin using the daily or periodical price rather than the market microstructure of this new financial asset.

Recent research concentrates on the microstructure of the Bitcoin market from the view point of order flow or order imbalance using historical trade data. For instance, Dimpfl (2017) inferred Bitcoin trade directions from tick data downloaded from bitcoinccharts.com. Feng et al. (2018) employed normalized order imbalance to measure informed trading in the Bitcoin market using microstructure data from bitcoinccharts.com. Wang et al. (2020) studied the impact of informed trading indexed by order imbalance on Bitcoin return and volatility using data from the same source as Dimpfl (2017) and Feng et al. (2018). Ibikunle et al. (2020) also utilized order imbalance in research on Bitcoin price discovery based on historical trade data from Bitstamp. On one hand, microstructure data with an initiator are not always available. It is important to assign trade directions based on classification algorithms when the initiator is not indicated in market microstructure research. As far as we know, the tick rule has been applied to classify historical trade data in Dimpfl (2017), Feng et al. (2018), Wang et al. (2020), and Ibikunle et al. (2020) for the Bitcoin market. On the other hand, the classification accuracy of the tick rule in the Bitcoin market, which consists of multiple online crypto exchanges, remains unknown.

This study investigates the accuracy of the tick rule in the Bitcoin market. The commonly used tick level data of Bitcoin, which are freely downloadable from bitcoinccharts.com, in research related to the Bitcoin market microstructure does not indicate the initiator. Nevertheless, the market data of Bitcoin/US Dollar (USD) with a trade indicator from Kaiko provide an opportunity to test the accuracy of the tick rule.

¹Shenzhen University, China
²Central South University, Changsha, China

Corresponding Author:
Pengxiang Zhai, Business School, Central South University, Changsha 410083, China.
Email: pengxiang.zhai@csu.edu.cn
rule in the Bitcoin market. This study focuses on three main questions: First, what is the classification success rate of the tick rule for Bitcoin market data? Second, what factors are associated with the classification accuracy of the tick rule in the Bitcoin market? Third, how accurate are the order imbalances based on the classification results of the tick rule in the Bitcoin market?

The primary results of this study are as follows. First, classification of 11.9 million Bitcoin/USD trades on Bitstamp suggests that the overall classification accuracy of the tick rule is 76.87%, which is close to that in the stock markets, and daily accuracy ranges from 68.98% to 83.76% during the sample period. This study finds a positive relationship between higher likelihood of misclassification and longer period between trades. As information spillover exists across Bitcoin exchanges (Brandvold et al., 2015), bigger time gaps contain more information to drive price changes. Furthermore, all of the order imbalances computed using the tick rule are significantly different from the true ones. The biases of the order imbalances calculated using large-size trades are relatively smaller than those using the whole sample.

The empirical findings of this study contribute to the small but ongoing exploration of the Bitcoin market microstructure in terms of research methodology. To the best of our knowledge, this is the first study to assess the classification accuracy of the tick rule in the Bitcoin market. Recently, tick-based classification methods have been used in research on the Bitcoin market microstructure (Dimpfl, 2017; Feng et al., 2018; Ibikunle et al., 2020; Wang et al., 2020) without knowledge of the classification accuracy of the tick rule. The conclusions from this study can thus serve as a guide for future research related to the methodological approach to be applied to the Bitcoin market microstructure when the transaction direction is not available.

This study also contributes to the existing literature on trade classification algorithms. Except for Aktas and Kryzanowski (2014) and Carrion and Kolay (2020), the classification accuracy of the tick rule in the above-mentioned articles is examined in slower trading environments, namely when high-frequency trading is not widely applied. Using tick data stamped to seconds from the Bitcoin market, this article also confirms the findings of Carrion and Kolay (2020), namely that the classification accuracy of the tick rule is similar to those in slower trading environments.

The remainder of this article is organized as follows. “Literature Review” section briefly reviews the literature on the tick rule. “Classification Accuracy” section reports on the empirical analysis of the classification accuracy of the tick rule. “Order Imbalance” section presents the biases of order imbalances based on trade directions assigned by the tick rule. “Discussion” section discusses the empirical findings of this study. Finally, “Conclusions” section concludes the article.

**Literature Review**

Trade classification approaches are commonly used to discern trade intentions when there is no available trade direction in market data. Popular classification methods include the tick rule or test, quote rule, Lee–Ready algorithm proposed by Lee and Ready (1991), and bulk volume classification (Easley et al., 2012, 2016). The tick rule assigns trade direction based on price movements when no quoted data are available. If the transaction price is higher than that of a previous transaction (uptick), then the present transaction is classified as a buyer-initiated order. Conversely, if the transaction price is below the latest price (downtick), then the present transaction is classified as a seller-initiated order. If there is no price change (zerotick), the transaction direction is assigned as the same as the previous one. On the other hand, the quote rule assigns trade as a buyer-initiated (seller-initiated) order if the price is above (below) the midpoint of the bid and the ask. Meanwhile, the Lee–Ready algorithm (Lee & Ready, 1991) combines the two aforementioned rules. For unclassified trades at midpoint, the Lee–Ready algorithm applies the tick rule to discern directions. Finally, bulk volume classification (Easley et al., 2012, 2016) uses empirical distribution of price changes to infer the possibility of buyer-initiated (seller-initiated) volume from the aggregate volume of each bar. As transaction direction is not always available when conducting researches on market microstructure, classification algorithms are very important to infer trade intentions in researches.

The tick rule is usually used in research related to market microstructure when only trade data are available. For example, Barber et al. (2009) employed the tick rule to identify directions of partial trades. Bernile et al. (2016) also used the tick rule to compute order imbalance, which measures informed trading activity ahead of the Federal Open Market Committee’s policy announcements. Dimpfl (2017), Feng et al. (2018), Wang et al. (2020), and Ibikunle et al. (2020) used the tick rule to assign directions to measure informed trading activity in the Bitcoin market.

Although the tick rule has been used in existing studies, the accuracy of such results is a cause for concern. Theissen (2001) noted that research related to the market microstructure model can be systematically biased because of inaccurate trade classification. Undoubtedly, the accuracy of the tick rule has been examined on stock markets. Using TORQ data for the U.S. stock market, Odders-White (2000) reported that 78.6% of transactions were correctly classified by the tick rule. Tests conducted by Finucane (2000) using information from the TORQ database from November 1990 to January 1991 showed that the classification accuracy of the tick rule applied to a sample of 144 NYSE firms was 83.0%. Ellis et al. (2000) utilized NASDAQ data during the period of September 27, 1996, to September 29, 1997, and showed that the success rate of the tick rule was 77.66%. Tests by
Chakrabarty et al. (2007) on NASDAQ stocks traded on INET and ArcaEx revealed that the overall success rate of the tick rule was 75.4% during the sample period. A recent examination by Carrion and Kolay (2020) showed that the classification success rate of the tick rule was 78.62% in a sample of data stamped to seconds from NASDAQ HFT database over a subset of dates during the period of 2008 to 2010 when trades were more frequent than before. And the classification success rate ranged from 69.75% to 83.34% across individual stocks. Similar studies have also focused on the non-U.S. stock markets. Aitken and Frino (1996) studied trades of 2 years on the Australian Stock Exchange and reported a success rate of 74% for the tick rule. Using 15 stocks on the Frankfurt Stock Exchange in 1996, Theissen (2001) documented that the tick rule correctly classified 72.2% of the transactions. An examination of classification algorithms on data from the Taiwan Stock Exchange by Lu and Wei (2009) revealed an overall success rate of 74.18% for the tick rule.Aktas and Kryzanowski (2014) examined the trade classification accuracy of different classification algorithms using the data of component firms of the BIST-30 index and found that the classification success rate of the tick rule ranged from 84.86% to 92.15% in different subsamples. In addition, Omrane and Welch (2016) found that the classification success rate of the tick rule for 1.2 million trades on the foreign exchange electronic communication network market was about 68%. To the best of the authors’ knowledge, none of the previous research has focused on the accuracy of trade classification algorithms for the cryptocurrency markets.

**Classification Accuracy**

**Data and Methodology**

Historical market data are required to test the accuracy of the tick rule in the Bitcoin market. Due to decentralization and lack of regulation, Bitcoin is traded simultaneously and continuously (24/7) on multiple online crypto exchanges. To examine the classification accuracy of the tick rule, the market data of Bitstamp, an order-driven online crypto exchange, is considered as being representative of its liquidity. Founded in Europe in 2011, Bitstamp is one of the oldest and largest global crypto exchanges that allows trading between Bitcoin and USD. The tick-by-tick trade data of Bitstamp were acquired from Kaiko, a cryptocurrency market data provider. Every transaction record includes a unique trade ID, timestamp, transaction price in USD, amount in Bitcoin, and trade direction indicator. This study includes classifiable trades from December 6, 2017, to October 7, 2018 (Greenwich Mean Time), amounting to a total of 11,919,298 observations. Since all these trades are stamped to seconds, some trades occur at the same timestamps because of fast trading. This study thus uses the trade ID to determine the order of the trades.

The detailed classification process of the tick rule applied to a specific tick $i$ can be summarized as seen in the following equation:

$$ Trade_i = \begin{cases} 
    \text{Buyer initiated}, & \text{if } P_i > P_{i-1} \\
    \text{Seller initiated}, & \text{if } P_i < P_{i-1} \\
    Trade_{i-1}, & \text{if } P_i = P_{i-1} 
\end{cases} $$

where $Trade_i$ is the estimated initiator of tick $i$, and $P_i$ is the executed price of tick $i$.

As defined in the previous literature, the classification accuracy $R_T$ during a specific period $T$ can be described by the following equation:

$$ R_T = \frac{n_r}{N_T} $$

where $N$ is the total number of transactions, and $n$ is the number of correctly classified transactions.

**Classification Accuracy**

Table 1 reports the classification accuracy of the tick rule for directions in Bitcoin transactions. According to the Bitstamp market data, true seller-initiated (buyer-initiated) orders account for 41.50% (58.50%) of all trades, whereas 45.84% (54.16%) of all trades are classified as seller-initiated (buyer-initiated) orders by the tick rule. Furthermore, the tick rule wrongly classifies 9.40% of all trades as buyer-initiated orders and 13.74% of all trades as seller-initiated orders. Thus, the overall classification success rate of the tick rule on the Bitstamp market data during the sample period is 76.87%.

Figure 1 shows the daily change in the misclassification rate. It can be observed that the misclassification rate of the tick rule varies with time, ranging from 16.24% to 31.02%. This means that daily success rate of the tick rule on the Bitcoin/USD transaction data of Bitstamp ranges from 68.98% to 83.76%.

| Direction   | Seller-initiated | Buyer-initiated | Sum     |
|-------------|------------------|-----------------|---------|
| Tick rule   | 3,826,489        | 1,637,398       | 5,463,887|
| Seller-initiated | (32.10%) | (13.74%) | (45.84%)|
| Tick rule   | 1,120,080        | 5,335,331       | 6,455,411|
| Buyer-initiated | (9.40%) | (44.76%) | (54.16%)|
| Sum         | 4,946,569        | 6,972,729       |         |
|             | (41.50%)         | (58.50%)        |         |

Note. This table reports the sample size of seller- and buyer-initiated transactions. The first column shows the number of seller-initiated transactions, including the number of ticks correctly classified as seller-initiated transactions and the number of ticks misclassified as buyer-initiated transactions. The second column presents the number of buyer-initiated transactions, including the number of ticks misclassified as seller-initiated transactions and the number of ticks correctly classified as buyer-initiated transactions. The numbers in the parentheses indicate the corresponding proportions of the total sample.
Table 2 displays the misclassification rate for each of the three tick types. The misclassification rates for uptick, downtick, and zerotick are 16.92%, 26.76%, and 25.28%, respectively, indicating that the classification success rate of the tick rule for upticks is much higher than that of the other tick types. In addition, the misclassification of zeroticks contributes the most to the total errors.

| Tick type | Total n | Total % | Misclassification n | Misclassification % |
|-----------|---------|---------|---------------------|---------------------|
| Uptick    | 3,662,689 30.73 | 619,572 16.92 |
| Downtick  | 3,412,012 28.63 | 913,135 26.76 |
| Zerotick  | 4,844,597 40.64 | 1,224,771 25.28 |

Note. This table reports the sample and misclassified sample sizes of different tick types. The first two columns report the tick number and proportion of sample size for each of the three tick types, whereas the last two columns provide the misclassified tick number and proportion in the corresponding tick type. Uptick is assigned when the transaction price is higher than that of a previous transaction. Conversely, downtick is assigned when the transaction price is below the latest price. In addition, zerotick is assigned when there is no price change.

Classification Under Different Market Conditions

Table 3 reports the classification accuracy of the tick rule during two subperiods to examine whether Bitcoin market conditions could affect the accuracy of this classification algorithm. First, Panel A displays the classification accuracy during the first subperiod from December 7, 2017, through March 31, 2018, which covered a highly bull market in December 2017 and a subsequent crash in January 2018 in Bitcoin price. It is found that the tick rule wrongly classifies 9.59% of all trades as buyer-initiated orders and 13.57% of all trades as seller-initiated orders. Namely, the classification accuracy of the tick rule is 76.84% during the first subperiod. Next, Panel B displays the classification accuracy of the second subperiod spanning from April 1, 2018, to October 7, 2018 during which Bitcoin price was relatively stable. Similarly, the classification accuracy of the tick rule is 76.90% during the second subperiod. Hence, Bitcoin market condition does not significantly impact the classification accuracy of the tick rule on the whole.

Multivariate Analysis

To analyze the variables associated with misclassification, this study draws from Ellis et al. (2000) and examines the following four variables: true trade direction, time from previous trade, trade size in Bitcoin, and price in USD. In addition, as trades are stamped to seconds, the time gaps of trades occurring at the same timestamp are set at zero.

Table 4 reports the distribution of misclassification rate of the tick rule in different subsamples. Panel A divides all trades into four groups according to the length of time from the previous trade. In the first group, where trades occur in no more
than 5 s, 22.19% of the trades are misclassified. Moreover, 33.14% of trades occurring in more than 60 s are misclassified. On the other hand, Panel B divides all trades into four groups based on trade size in Bitcoin. The classification success rate is the largest when the order amount is no greater than 0.01 Bitcoin among the four subsamples. However, no monotonic relation is observed between trade size and classification success. Last, Panel C divides all trades into four groups based on executed price in USD. Notably, 21.57% of trades in the highest price group are misclassified, while the misclassification rate in the lowest price group is 23.79%. Considering that all these variables originated from the same

Figure 2. Proportion of seller-initiated orders: (A) daily change of misclassified sell/misclassified sample; (B) daily change of sell/total.

Table 3. Classification Accuracy During Subperiods.

| Direction                  | Seller-initiated | Buyer-initiated | Sum          |
|----------------------------|------------------|----------------|--------------|
|                            |                  |                |              |
| Panel A: From December 7, 2017, to March 31, 2018 |                  |                |              |
| Tick rule                  | 2,322,029        | 939,273        | 3,261,302    |
| Seller-initiated           | (33.55%)         | (13.57%)       | (47.12%)     |
| Tick rule                  | 664,017          | 2,996,583      | 3,660,600    |
| Buyer-initiated            | (9.59%)          | (43.29%)       | (52.88%)     |
| Panel B: From April 1, 2018, to October 6, 2018 |                  |                |              |
| Tick rule                  | 1,504,460        | 698,125        | 2,202,585    |
| Seller-initiated           | (30.10%)         | (13.97%)       | (44.07%)     |
| Tick rule                  | 456,063          | 2,338,748      | 2,794,811    |
| Buyer-initiated            | (9.13%)          | (46.80%)       | (55.93%)     |

Note. This table reports the sample size of seller- and buyer-initiated transactions during each subperiod. The first column shows the number of seller-initiated transactions, including the number of ticks correctly classified as seller-initiated transactions and the number of ticks misclassified as buyer-initiated transactions. The second column presents the number of buyer-initiated transactions, including the number of ticks misclassified as seller-initiated transactions and the number of ticks correctly classified as buyer-initiated transactions. The numbers in the parentheses indicate the corresponding proportions of the total sample during each subperiod.
data, correlations may exist among them. Further studies on the relationship between misclassification and time from previous trade/trade size/price level are thus needed.

Table 5 reports the results of multivariate regressions, including ordinary least square (OLS) and logistic regressions. At the beginning, multivariate regressions of all trades indicate that seller-initiated order, amount, and price are negatively associated with the likelihood of misclassification, while time from previous trade is positively correlated with the likelihood of misclassification. However, the regression results of the subsamples are not in line with those of the whole sample. The results show that when time from the previous trade is no longer than 5 s, the likelihood of misclassification is positively associated with seller-initiated order and time from the previous trade, and negatively associated with trade size and price level. When trade size is no more
than 0.1 Bitcoin, the likelihood of misclassification increases with all four independent variables. However, for an executed price of no more than 10,000 USD, the likelihood of misclassification decreases with price. In general, it is evident that the likelihood of misclassification is positively associated with time between trades in all regressions.

**Order Imbalance**

Order imbalance is usually employed as a measure of informed trading activity that cannot be observed directly. Order imbalance here is defined as \((B - S)/(B + S)\), where \(B\) (\(S\)) denotes buyer-initiated (seller-initiated) variables. This study examines imbalances (Barber & Odean, 2008; Bernile et al., 2016; Feng et al., 2018; Ibikunle et al., 2020; Ning & Tse, 2009; Sun & Ibikunle, 2017; Wang et al., 2020) estimated using the tick rule, namely order imbalance based on the number of trades (OIN), trade size (OIS), and volume in USD (OID), using market data from Bitstamp. In addition, large-size orders are usually considered to be related to informed trading, because informed traders are prone to using block trades to cut down transaction costs. Following Feng et al. (2018) and Wang et al. (2020), the order imbalances of large-size orders that are larger than the 95th percentile of intraday trade sizes are also compared for robustness.

Figures 3 to 5 present daily order imbalances measured by OIN, OIS, and OID, respectively. Plots A and C in Figure 3 are true OINs calculated with the number of true trade directions and OINs calculated with directions assigned by the tick rule, respectively. Plot E in Figure 3 shows the bias of the estimated OIN, defined as the difference between the true OIN and the OIN estimated using the tick rule. Plots B, D, and F in Figure 3 use the same method but for large-size trades. The two lines in Plots E and F indicate the values of \(-0.1\) and \(0.1\). The proportions of OIN bias whose absolute value is larger than 0.1 are 148/306≈48.37% in Plot E and 74/306≈24.18% in Plot F. Figures 4 and 5 report the daily order imbalances measured by OIS and OID, respectively, using the same method. Therefore, the proportions of biases whose absolute value is larger than 0.1 are 70/306≈22.88% (OIS) and 73/306≈23.86% (OID) in Figure 4, and 69/306≈22.55% (OID) and 72/306≈23.53% (OID) in Figure 5. In general, all these order imbalance measures are biased to a certain degree.

Table 6 reports the results of a parametric test (the Welch two-sample \(t\) test) and a nonparametric test (the Mann– Whitney \(U\) test) applied to examine the differences between the true and estimated order imbalances. On one hand, the results of the Welch two-sample \(t\) test reveal that order imbalances estimated using the tick rule are underestimated on the whole and are statistically different from the true order imbalances at the 5% level. On the other hand, the null hypotheses of the Mann–Whitney \(U\) test are all rejected at the 1% significance level, indicating that distributions of the true and estimated order imbalances are not equal. On the whole, the statistics in Table 6 suggest that order imbalances in the Bitcoin market computed using the tick rule lack adequate accuracy.

Table 7 reports the regression results of daily return and volatility of Bitcoin on each daily order imbalance to explore whether order imbalances could predict Bitcoin return or volatility. At the beginning, the first three columns display estimates of order imbalances from regressions of Bitcoin daily return. And it is found that all order imbalances (including the true and estimated ones marked by “TR”) are positively correlated with daily return at the 1% significance level. However, the adjusted R-squared shows that order imbalances computed using the tick rule rather than the true ones could predict more variation of Bitcoin daily return. Then, the next three columns display estimates of order imbalances from regressions of Bitcoin realized variance (RV, hereafter) multiplied by \(10^4\). Commonly, RV proposed by Andersen and Bollerslev (1998) is used as an ex-post volatility measure in financial literature. In this study, all estimates of order imbalances are negative and statistically significant at the 5% level, which means that realized variance decreases as order imbalances increase. Nevertheless, the adjusted R-squared of each order imbalance is less than 10% whereas order imbalances based on number of trades including OIN and OIN (TR) outperform others. Third, the rest of Table 7 displays estimates of order imbalances from regressions of positive semi-variance \(RV^+_d\) multiplied by \(10^4\) and negative semi-variance \(RV^-_d\) multiplied by \(10^4\). And it is found that order imbalances are more statistically correlated with negative semi-variance \(RV^-_d\), and predict more variation of negative semi-variance \(RV^-_d\) than that of positive semi-variance \(RV^+_d\). On the whole, it is not very surprising that order imbalances computed using the tick rule outperform the true ones in predicting daily return because they are ex-post measures based on price change or return. In other words, performance of true order imbalance is overestimated by using estimated order imbalance in predicting daily return during sample period.

**Discussion**

Given that Bitcoin is listed on multiple unregulated online crypto exchanges, the classification accuracy of the tick rule in the Bitcoin market is similar to those in stock markets. The empirical analysis in this study shows that the overall classification accuracy is 76.87% and the daily classification accuracy ranges from 68.98% to 83.76% in the Bitcoin market. According to previous research, the classification success rate of the tick rule ranges from 72.2% (Theissen, 2001) to 92.15% (Aktas & Kryzanowski, 2014) on the U.S. and non-U.S. stock markets. Of the research cited in this work, the study by Carrion and Kolay (2020) presents the similar fast trading environment by using high-frequency NASDAQ data stamped to seconds. And the accuracy of the tick rule
assessed in this study is close to the corresponding values of individual stocks in Carrion and Kolay (2020), namely from 69.75% to 83.34%.

Therefore, the empirical results indicate that there exists a positive correlation between the likelihood of misclassification and the time from the previous trade in the Bitcoin market, as shown in Tables 4 and 5. Conversely, Ellis et al. (2000) found a higher classification success rate when trades were slow, due to a higher turnover rate of quotes. The difference can be attributed to the fact that since Bitcoin is traded simultaneously on multiple online crypto exchanges, information spillover from other crypto exchanges could impact the price (Brandvold et al., 2015). Consequently, it may be more difficult to discern trade direction based on previous trade when a long period of time elapses between trades.

Figure 3. Order imbalance based on number of trades (OIN): (A) true OIN; (B) true OIN95; (C) OIN estimated using tick rule; (D) OIN95 estimated using tick rule; (E) bias of OIN; (F) bias of OIN95.

Note. OIN is defined as \( \frac{B_n - S_n}{B_n + S_n} \), where \( B_n \) and \( S_n \) are the counts of buyer- and seller-initiated trades, respectively. OIN95 is calculated using trades whose sizes in Bitcoin are larger than the 95th percentile of intraday trade sizes. The two lines in Plots E and F indicate the values of −0.1 and 0.1.
In addition, order imbalances calculated using large-size trades are relatively closer to their true values in the Bitcoin market during the sample period. As shown in Table 6, the results of the Welch two-sample *t* test show that the differences in the means of the true and estimated order imbalances are smaller and less statistically significant when the order imbalances are calculated using trade sizes larger than the 95th percentile of intraday trade sizes.

**Conclusions**

This study investigates the accuracy of the tick rule in the Bitcoin market, wherein Bitcoin is listed on multiple online crypto exchanges rather than traditional regular exchanges. Although the tick rule has been utilized in researches on the microstructure of this innovational market (Dimpfl, 2017; Feng et al., 2018; Ibikunle et al., 2020; Wang et al., 2020),
Figure 5. Order imbalance based on volume in USD (OID): (A) true OID; (B) true OID95; (C) OID estimated using tick rule; (D) OID95 estimated using tick rule; (E) bias of OID; (F) bias of OID95.

Note. OID is defined as \( \frac{BD - SD}{BD + SD} \), where \( BD \) and \( SD \) are the volumes of buyer- and seller-initiated trades in USD, respectively. OID95 is calculated using trades whose sizes in Bitcoin are larger than the 95th percentile of intraday trade sizes. The two lines in Plots E and F indicate the values of \( -0.1 \) and \( 0.1 \).

Table 6. Difference Between True and Estimated Order Imbalances.

| OI     | Welch two-sample t test | Mann–Whitney U test |
|--------|-------------------------|---------------------|
|        | Mean of true (%)        | Mean of estimated (%)| t-value | Statistic     |
| OIN    | 19.08                   | 9.91                | 9.44*** | 66,172***     |
| OIN95  | 2.02                    | −2.17               | 4.40*** | 56,888***     |
| OIS    | 2.18                    | −2.15               | 5.37*** | 58,333***     |
| OIS95  | 1.19                    | −1.87               | 2.65**  | 52,864**      |
| OID    | 2.25                    | −2.01               | 5.30*** | 58,180***     |
| OID95  | 1.29                    | −1.69               | 2.58*   | 52,708*       |

Note. OI = order imbalance; OIN = order imbalance based on number of trades; OIS = order imbalance based on trade size; OID = order imbalance based on volume in USD.

*, **, and *** denote statistical significance at the 5%, 1%, and 0.1% levels, respectively.
Table 7. Predictability of Order Imbalance.

| OI | Estimate | t-value | Adj-R² | Estimate | t-value | Adj-R² | Estimate | t-value | Adj-R² |
|----|----------|---------|--------|----------|---------|--------|----------|---------|--------|
| OIN | 7.42e-2 | 3.68 | 4.18% | -113.81** | -3.13 | 7.41% | -59.38** | -2.96 | 7.08% |
| OIN(TR) | 1.91e-1 | 7.35 | 14.07% | -181.19*** | -3.52 | 9.17% | -86.11*** | -3.26 | 7.21% |
| OIN95 | 1.61e-1 | 5.09 | 14.56% | -91.63*** | -2.82 | 3.19% | -43.57* | -2.52 | 2.47% |
| OIN95(TR) | 2.01e-1 | 5.53 | 23.68% | -63.60** | -2.67 | 1.42% | -23.52 | -1.95 | 0.51% |
| OIS | 2.27e-1 | 7.04 | 21.14% | -95.51** | -2.64 | 2.45% | -45.25* | -2.28 | 1.87% |
| OIS(TR) | 2.91e-1 | 7.07 | 35.02% | -66.57* | -2.35 | 1.03% | -21.36 | -1.49 | 0.16% |
| OIS95 | 1.63e-1 | 5.21 | 21.14% | -67.37** | -2.84 | 2.34% | -29.64* | -2.35 | 0.51% |
| OIS95(TR) | 1.80e-1 | 6.47 | 29.03% | -46.94** | -2.75 | 1.13% | -15.81 | -1.49 | 0.25% |
| OID | 2.27e-1 | 7.01 | 20.98% | -91.15* | -2.53 | 2.19% | -42.91* | -2.18 | 1.64% |
| OID(TR) | 2.92e-1 | 7.05 | 35.18% | -61.59** | -2.18 | 0.81% | -18.27 | -1.27 | 0.03% |
| OID95 | 1.63e-1 | 5.19 | 21.06% | -64.68** | -2.81 | 2.14% | -28.19* | -2.31 | 1.32% |
| OID95(TR) | 1.81e-1 | 6.45 | 29.17% | -43.74* | -2.59 | 0.93% | -13.98 | -1.59 | 0.12% |
| OI95 | 2.27e-1 | 7.04 | 20.98% | -91.15* | -2.53 | 2.19% | -42.91* | -2.18 | 1.64% |
| OI95(TR) | 2.92e-1 | 7.05 | 35.18% | -61.59** | -2.18 | 0.81% | -18.27 | -1.27 | 0.03% |

Note. This table reports the regression results whereas estimates of intercepts are not displayed to enhance readability. The model can be expressed as $y_i = \alpha + \beta x_i + \epsilon_i$ where the dependent variable $y_i$ is the daily return or volatility of Bitcoin on day $d$; the independent variable $x_i$ is the daily order imbalance; $\alpha$ is the intercept; and $\epsilon_i$ denotes the error item. Then, return in this study is defined as $R_t = P_t / P_{t-1} - 1$ where $P_t$ is closing price at time $t$. Following Eross et al. (2019), we use the RV computed with 5 min high-frequency data aggregated from the tick data to capture the volatility of Bitcoin return. RV of day $d$ is equal to $RV_d = \sum_{i=1}^{T} \epsilon_i^2$. Furthermore, consistent with Barndorff-Nielsen et al. (2010), we decompose $RV_d$ into two parts, i.e., positive semi-variance $RV_d^+$ and negative semi-variance $RV_d^-$. $RV_d^-$ and $RV_d^+$ can be defined as $RV_d^+ = \sum_{i=1}^{T} \epsilon_i^2, I(\epsilon_i > 0)$ and $RV_d^- = \sum_{i=1}^{T} \epsilon_i^2, I(\epsilon_i < 0)$ where $RV_d = RV_d^+ + RV_d^-$. All t-values reported here are adjusted for Newey-West standard errors. RV = realized variance; OI = order imbalance; OIN = order imbalance based on number of trades; OIS = order imbalance based on trade size; OID = order imbalance based on volume in USD; TR = tick rule. * , **, and *** denote statistical significance at the 5%, 1%, and 0.1% levels, respectively.

the accuracy of this trade classification method requires further examination. This study attempts to address three issues: the success rate of the tick rule in the Bitcoin market, factor(s) associated with classification success, and bias of order imbalances, which are usually used as indexes for informed trading computed using the tick rule.

This study answers the three above-stated questions through empirical analysis using the tick-by-tick transaction data of Bitcoin/USD with signed initiators on Bitstamp from December 6, 2017, to October 7, 2018. First, this study finds that the overall success rate of the tick rule is 76.87%, and the daily accuracy ranges from 68.98% to 83.76% during the sample period. There are less misclassified seller-initiated orders than misclassified buyer-initiated ones on the whole, and this result is associated with fewer seller-initiated trades in the sample. In general, trade classification using the tick rule in the Bitcoin market has limited success. Second, this study finds that the longer the time between trades, the higher the possibility of misclassification. It is more difficult to discern transaction intentions when transactions are less frequent in this innovative market of multiple online crypto exchanges. Third, the empirical analysis indicates that the order imbalances computed using the tick rule in the Bitcoin market lack sufficient accuracy. However, order imbalances calculated using large-size trades are relatively closer to their true values. Evidently, attention must be paid to the accuracy of the trade classification algorithm when conducting research on the microstructure of the Bitcoin market.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by the Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (C) 17K03657.

ORCID iD
Donglian Ma [ID] https://orcid.org/0000-0002-8885-0477

References
Aitken, M., & Frino, A. (1996). The accuracy of the tick test: Evidence from the Australian stock exchange. *Journal of Banking & Finance*, 20(10), 1715–1729. https://doi.org/10.1016/S0378-4266(96)00008-8
Aktaş, O. U., & Kryzanowski, L. (2014). Trade classification accuracy for the BIST. *Journal of International Financial Markets, Institutions and Money*, 33, 259–282. https://doi.org/10.1016/j.jfinm.2014.08.003
Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts.
Andrade-Félix, J., Fernandez-Perez, A., & Sosvilla-Rivero, S. (2020). Distant or close cousins: Connectedness between cryptocurrencies and traditional currencies volatilities. Journal of International Financial Markets, Institutions and Money, 67, Article 101219. https://doi.org/10.1016/j.jifm.2020.101219

Ante, L., & Fiedler, I. (2020). Market reaction to large transfers on the Bitcoin blockchain—Do size and motive matter? Finance Research Letters, 39, Article 101619. https://doi.org/10.1016/j.frl.2020.101619

Antonakakis, N., Chatziantoniou, I., & Gabauer, D. (2019). Cryptocurrency market contagion: Market uncertainty, market complexity, and dynamic portfolios. Journal of International Financial Markets, Institutions and Money, 61, 37–51. https://doi.org/10.1016/j.jifm.2019.02.003

Baek, C., & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. Applied Economics Letters, 22(1), 30–34. https://doi.org/10.1080/13504851.2014.916379

Baig, A., Blau, B. M., & Sabah, N. (2019). Price clustering and sentiment in Bitcoin. Finance Research Letters, 29, 111–116. https://doi.org/10.1016/j.frl.2019.03.013

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2), 785–818. https://doi.org/10.1093/rfs/hhn079

Barber, B. M., Odean, T., & Zhu, N. (2009). Do retail trades move markets? The Review of Financial Studies, 22(1), 151–186.

Barviera, A. F. (2017). The inefficiency of Bitcoin revisited: A dynamic approach. Economics Letters, 161, 1–4.

Barviera, A. F., Zunino, L., & Rosso, O. A. (2018). An analysis of high-frequency cryptocurrencies prices dynamics using permutation-information-theory quantifiers. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28(7), Article 075511. https://doi.org/10.1063/1.5027153

Barndorff-Nielsen, O. E., Kinnebrock, S., & Shephard, N. (2010). Cryptocurrency market contagion: Market uncertainty, market complexity, and dynamic portfolios. Journal of International Financial Markets, Institutions and Money, 61, 37–51. https://doi.org/10.1016/j.jifm.2019.02.003

Baur, C., & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. Applied Economics Letters, 22(1), 30–34. https://doi.org/10.1080/13504851.2014.916379

Baig, A., Blau, B. M., & Sabah, N. (2019). Price clustering and sentiment in Bitcoin. Finance Research Letters, 29, 111–116. https://doi.org/10.1016/j.frl.2019.03.013

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of Financial Studies, 21(2), 785–818. https://doi.org/10.1093/rfs/hhn079

Barber, B. M., Odean, T., & Zhu, N. (2009). Do retail trades move markets? The Review of Financial Studies, 22(1), 151–186.

Barviera, A. F. (2017). The inefficiency of Bitcoin revisited: A dynamic approach. Economics Letters, 161, 1–4.

Barviera, A. F., Zunino, L., & Rosso, O. A. (2018). An analysis of high-frequency cryptocurrencies prices dynamics using permutation-information-theory quantifiers. Chaos: An Interdisciplinary Journal of Nonlinear Science, 28(7), Article 075511. https://doi.org/10.1063/1.5027153

Brandvold, M., Molnár, P., Vagstad, K., & Valstad, O. C. A. (2015). Price discovery on Bitcoin exchanges. Journal of International Financial Markets, Institutions and Money, 36, 18–35. https://doi.org/10.1016/j.jifm.2015.02.010

Carrión, A., & Kolay, M. (2020). Trade signing in fast markets. Financial Review, 55(3), 385–404. https://doi.org/10.1111/fire.12218

Chakrabarty, B., Li, B., Nguyen, V., & Van Ness, R. A. (2007). Trade classification algorithms for electronic communications network trades. Journal of Banking & Finance, 31(12), 3806–3821. https://doi.org/10.1016/j.jbankfin.2007.03.003

Corbet, S., Lucey, B., & Yarovaya, L. (2018). Datestamping the Bitcoin and Ethereum bubbles. Finance Research Letters, 26, 81–88. https://doi.org/10.1016/j.frl.2017.12.006

Dimpfl, T. (2017). Bitcoin market microstructure. https://doi.org/10.2139/ssrn.2949807

Dyhrberg, A. H. (2016a). Bitcoin, gold and the dollar—A GARCH volatility analysis. Finance Research Letters, 16, 85–92. https://doi.org/10.1016/j.frl.2015.10.008

Dyhrberg, A. H. (2016b). Hedging capabilities of Bitcoin. Is it the virtual gold? Finance Research Letters, 16, 139–144. https://doi.org/10.1016/j.frl.2015.10.025

Easley, D., de Prado, M. L., & O’Hara, M. (2016). Discerning information from trade data. Journal of Financial Economics, 120(2), 269–285. https://doi.org/10.1016/j.jfineco.2016.01.018

Easley, D., López de Prado, M. M., & O’Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. The Review of Financial Studies, 25(5), 1457–1493. https://doi.org/10.1093/rfs/hls053

Ellis, K., Michaely, R., & O’Hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. The Journal of Financial and Quantitative Analysis, 35(4), 529–551. https://doi.org/10.2307/2676254

Eross, A., McGroarty, F., Urquhart, A., & Wolfe, S. (2019). The intraday dynamics of Bitcoin. Research in International Business and Finance, 49, 71–81. https://doi.org/10.1016/j.ribaf.2019.01.008

Feng, W., Wang, Y., & Zhang, Z. (2018). Informed trading in the Bitcoin market. Finance Research Letters, 26, 63–70. https://doi.org/10.1016/j.frl.2017.11.009

Finucane, T. J. (2000). A direct test of methods for inferring trade direction from intra-day data. The Journal of Financial and Quantitative Analysis, 35(4), 553–576. https://doi.org/10.2307/2676255

Hu, B., McInish, T., Miller, J., & Zeng, L. (2019). Intraday price behavior of cryptocurrencies. Finance Research Letters, 28, 337–342. https://doi.org/10.1016/j.frl.2018.06.002

Ibikunle, G., McGroarty, F., & Rzayev, K. (2020). More heat than light: Investor attention and Bitcoin price discovery. International Review of Financial Analysis, 69, Article 101459. https://doi.org/10.1016/j.irf.2020.101459
Klein, T., Thu, H. P., & Walther, T. (2018). Bitcoin is not the new gold—A comparison of volatility, correlation, and portfolio performance. *International Review of Financial Analysis, 59*, 105–116. https://doi.org/10.1016/j.irfa.2018.07.010

Koutmos, D. (2018a). Bitcoin returns and transaction activity. *Economics Letters, 167*, 81–85. https://doi.org/10.1016/j.econlet.2018.08.041

Kurihara, Y., & Fukushima, A. (2017). The market efficiency of Bitcoin: A weekly anomaly perspective. *Journal of Applied Finance & Banking, 7*(3), 57–64.

Kurka, J. (2019). Do cryptocurrencies and traditional asset classes influence each other? *Finance Research Letters, 31*, 38–46. https://doi.org/10.1016/j.frl.2019.04.018

Kwon, J. H. (2020). Tail behavior of Bitcoin, the dollar, gold and the stock market index. *Journal of International Financial Markets, Institutions and Money, 67*, Article 101202. https://doi.org/10.1016/j.intfin.2020.10.005

Lee, C. M., & Ready, M. J. (1991). Inferring trade direction from intraday data. *The Journal of Finance, 46*(2), 733–746. https://doi.org/10.1111/j.1540-6261.1991.tb02683.x

Li, X., Li, S., & Xu, C. (2020). Price clustering in Bitcoin market—An extension. *Finance Research Letters, 32*, Article 101072. https://doi.org/10.1016/j.frl.2018.12.020

Lu, Y. C., & Wei, Y. C. (2009). Classification of trade direction for an equity market with price limit and order match: Evidence from the Taiwan stock market. *Investment Management and Financial Innovations, 6*(3), 135–147.

Nadarajah, S., & Chu, J. (2017). On the inefficiency of Bitcoin. *Economics Letters, 150*, 6–9. https://doi.org/10.1016/j.econlet.2016.10.033

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*. https://bitcoin.org/en/

Nguyen, L. H., Chevapatrakul, T., & Yao, K. (2020). Investigating tail-risk dependence in the cryptocurrency markets: A LASSO quantile regression approach. *Journal of Empirical Finance, 58*, 333–355. https://doi.org/10.1016/j.jempfin.2020.06.006

Ning, Z., & Tse, Y. (2009). Order imbalance in the FTSE index futures market: Electronic versus open outcry trading. *Journal of Business Finance and Accounting, 36*(1–2), 230–252. https://doi.org/10.1111/j.1468-5957.2008.02116.x

Odders-White, E. R. (2000). On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets, 3*(3), 259–286. https://doi.org/10.1016/S1386-4181(00)00006-9

Omran, W. B., & Welch, R. (2016). Tick test accuracy in foreign exchange ECN markets. *Research in International Business and Finance, 37*, 135–152. https://doi.org/10.1016/j.ribaf.2015.10.001

Sun, Y., & Ibikunle, G. (2017). Informed trading and the price impact of block trades: A high frequency trading analysis. *International Review of Financial Analysis, 54*, 114–129. https://doi.org/10.1016/j.irfa.2016.07.005

Theissen, E. (2001). A test of the accuracy of the Lee/Ready trade classification algorithm. *Journal of International Financial Markets, Institutions and Money, 11*(2), 147–165. https://doi.org/10.1016/S1042-4431(00)00048-2

Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters, 150*, 1–7. https://doi.org/10.1016/j.econlet.2016.09.019

Wang, J. N., Liu, H. C., Zhang, S., & Hsu, Y. T. (2020). How does the informed trading impact Bitcoin returns and volatility? *Applied Economics, Advance online publication*. https://doi.org/10.1080/00036846.2020.1814944

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. In D. L. K. Chuen (Ed.), *Handbook of digital currency* (pp. 31–43). Academic Press. https://doi.org/10.1016/B978-0-12-802117-0.00002-3

Zargar, F. N., & Kumar, D. (2019). Informational inefficiency of Bitcoin: A study based on high-frequency data. *Research in International Business and Finance, 47*, 344–353. https://doi.org/10.1016/j.ribaf.2018.08.008