ESG UQÀM

Centre d'intelligence en surveillance des marchés financiers

Les cahiers de recherche du CISMF
CISMF Research Paper Series

Order-Driven Markets: Spoofing the Spoofers

Michèle Breton [a] & Amin Nejat[b]

[a] Department of Decision Sciences, HEC Montréal, Canada
[b] Department of Applied Economics, HEC Montréal, Canada

Février / February 2024

Centre d'intelligence en surveillance des marchés financiers

Order-Driven Markets: Spoofing the Spoofers

Amin Nejat[1], Michèle Breton[2]
[1]Department of Applied Economics, HEC Montreal – Quebec, Canada
[2]Department of Decision Sciences, HEC Montreal – Quebec, Canada

February 16, 2024

Abstract

Spoofing is one of the forms of market manipulation, which consists of placing orders with the intent to cancel them before execution, in order to artificially influence the market price and mislead investors. This paper uses cryptocurrency market data to propose price-impact models, which are the foundation of market manipulation strategies. In addition, we show how order-book data could be used to detect and/or prevent manipulation in order-driven markets.

1 INTRODUCTION

Market manipulation is the act of artificially raising or lowering the price of a security so that its price differs from its true value, or otherwise influencing the behavior of the market for personal gain (https : //www.investopedia.com/terms/m/manipulation.asp). The goal of market manipulation is to deceive other market participants in order to create a situation where assets are mispriced, so that the manipulator (who knows better) can then profit from the situation, at the expense of other market participants. Because the manipulators themselves create (and subsequently reverse) the mispricing in the first place, market manipulation, unlike honest investing strategies, does not improve market efficiency or benefit society. On the contrary, market manipulation is detrimental to the well-functioning of financial markets, as it distorts prices, interferes with the efficient allocation of resources, and may deter investors from participating.

Market manipulation can take many forms, but the most common methods are based on either fake orders, fake news or statements, or on artificial activity. Manipulation is illegal in most cases, but it may be difficult to detect and prosecute for regulators and other authorities.

Cryptocurrency is a relatively new class of digital asset that has recently become very popular. However, cryptocurrencies seem to be a fertile field for price manipulation, due to a number of specific features, such as lack of regulation, non-stop trading in multiple exchanges, high volumes and volatility, and automated trading.

In this paper, we investigate spoofing, which is a form of market manipulation commonly used in cryptocurrency markets. The Dodd-Frank Act (see Public Law 111-203 2010) describes spoofing as "bidding or offering with the intent to cancel the bid or offer before execution".

Spoofing (also called layering) consists of placing a relatively large number of bids to buy (*resp. offers to sell*) assets, and canceling the bids (*resp. offers*) prior to execution. Manipulators spoof the market, creating a false picture of either high demand or of pessimism, by posting fake quotes on one side of the limit order book. Spoofing may cause prices to change because the market interprets the one-sided pressure in the limit order book as a shift in the balance of the number of investors who wish to purchase or sell the asset, which causes prices to increase (more buyers than sellers) or to decline (more sellers than buyers). The flurry of activity around the fake orders is intended to attract other traders and to induce them to trade in anticipation of expected price movements. Spoofing is difficult to detect, as the manipulator's activities blend with the huge number of updates in the order book and are produced using sophisticated automated algorithms to avoid detection.

2 Review of related literature

While there exists an extensive theoretical literature regarding manipulative stock trading (see for instance Allen & Gale 1992), empirical research on spoofing is relatively scarce.

Lee et al. (2013) defines a spoof order as a quote at least 6 ticks away from the market price, with a volume at least twice as large as the average volume of the orders posted on the previous day. The authors use a proprietary data set with account information from the Korea Exchange, and show empirically that the spoof orders create imbalance in the order book, which moves the price. The authors also show empirically that spoofing achieves substantial extra profits and tends to target stocks with higher volatility of returns, lower market capitalization, and lower price level.

Wang (2015) uses data from the Taiwan Futures Exchange and shows that spoofers manipulate the order book of stocks that have high volumes of trading, high volatility, and high prices. The author's findings also suggest that spoofing increases trading volume, asset price volatility and bid-ask spread.

Cartea et al. (2020) derives an optimal trading strategy for spoofers, accounting for the trade-off between the benefits from spoofing and the potential fine levied from the manipulator by the financial authorities. The authors show that spoofing does deviate the price of the asset from its fundamental value. As expected, this deviation is highest when market participants believe the information conveyed by the order book, and the fine for spoofing is null.

In the particular case of cryptocurrencies, Gandal et al. (2018) and Griffin & Shams (2020) analyze possible episodes of market manipulations of the Bitcoin price. Twomey & Mann (2020) and Eigelshoven et al. (2021) identify and classify various cryptocurrency market manipulation schemes. Other papers examine specific manipulation schemes, such as wash trading (Aloosh & Li 2019, Le Pennec et al 2021, Cong et al 2023) and "pump & dump" (Chen et al 2019, Xu et al 2019, Mansourifar et al 2020, Hamrick et al 2021).

As in Cao et al (2014), we propose the use of order book data to detect price manipulation. We specifically look for spoofing and use Bitcoin market data. Since spoofing aims at influencing the asset price through order imbalance, we propose in the next section a price-impact model involving volumes of the bid and ask side in the limit order book.

3 Price-impact model

3.1 Data

Our data set comprises limit order book data and trades attributes from the Coinbase exchange between January 1st 2021 and March 2nd 2021. Since Bitcoin is a 24/7 market, the data is not limited to any time interval, contrary to traditional markets. The data is collected through the API provided by Coinbase.

Table 1 provides the description of the variables collected from the Coinbase database.

3.2 Variables

The predictors used in the literature for the price impact function usually include volume and volatility. The set of variables for quote and trade prices and volumes

Table 1: Definition of variables collected through Coinbase API.

Variables	Description
$\nu^a_{\tau i}$, $\nu^b_{\tau i}$	Volume of i^{th} quote in the ask (resp. bid) side of order book at time τ
$p^a_{\tau i}$, $p^b_{\tau i}$	Price of i^{th} quote in the ask (resp. bid) side of order book at time τ
ν^B_τ, ν^S_τ	Size of a buy (resp. sell) initiated trade fulfilled at time τ
p^B_τ, p^S_τ	Size of a buy (resp. sell) initiated trade fulfilled at time τ

extracted from the collected data and used in our regression model is given in Table 2 and defined precisely below.

Table 2: Definition of features used in the regression model.

Feature	Description
Δ_t	Change in mean of mid-price from time interval t to $t+1$
V^O_t	Difference between bid side and ask side quotes' volume in interval t
P^O_t	Sum of ask and bid quotes' price dispersion from mid-price in interval τ
V^T_t	Difference between buy-initiated and sell-initiated trades' volume in interval τ
D_{tq}	Distance of mid-price from the moving average price calculated at t for last q intervals
r_{tq}	Ratio of buy or sell initiated trades volume at t to all corresponding trades for the last q intervals
σ_{tq}	Volatility of Bitcoin price return at t for the last q intervals
δ_t	Indicator of the sign of Δ_t

Δ_t is the dependent variable, measuring the change in mean of mid-price from time interval t to $t+1$:

$$\Delta_t = \frac{\overline{P}^m_{t+1}}{\overline{P}^m_t}$$

where $\overline{P}^m_t$ is the average of observed mid prices p^m_τ during the time interval t, the mid price p^m_τ at τ is the average of the best bid and ask quote prices,

$$p^m_\tau = \frac{\max_i p^b_{\tau i} + \min_i p^a_{\tau i}}{2},$$

and where τ is an instant during t-th interval.

The volume variable V^O_t measures the difference in volumes of the two sides of the order book during the time interval t

$$V^O_t = \frac{\sum_{\tau=1}^{n_t} \sum_{i=1}^{m_\tau} \nu^b_{\tau i} - \nu^a_{\tau i}}{n_t},$$

where n_t is the number of instant observations during the interval t and m_τ is the number of quotes at time τ.

The price dispersion variable P_t^O during time interval t is computed using

$$P_t^O = \frac{\sum_{\tau=1}^{n_t} \sum_{i=1}^{m_\tau} \left(\frac{p_{\tau i}^b}{p_\tau^m} - 1 \right) + \left(\frac{p_{\tau i}^a}{p_\tau^m} - 1 \right)}{n_t},$$

where $p_{\tau i}^b$ (*resp.* $p_{\tau i}^a$) is the price of the i-th quote of the bid (*resp. ask*) side of the limit order book at time τ, and p_τ^m is the mid-price at τ. This variable measures the dispersion of the quotes from the mid price during a given time interval.

The difference variable V_t^T is computed using

$$V_t^T = \sum_{\tau=1}^{n_t} \nu_\tau^B - \nu_\tau^S,$$

where ν_τ^B and ν_τ^S are the volume of buy and sell trades fulfilled at date τ.

The distance variable D_{tq} is computed using

$$D_{tq} = \frac{\overline{P}_t^m}{\frac{1}{q} \sum_{q=0}^{l} \overline{P}_{t-q}^m} - 1,$$

where the denominator represents the moving average of prices in the q-interval period prior to t.

The initiated trade volumes (buy or sell, depending on which is the largest) during time interval t is compared to the average over the q preceding intervals using

$$r_{tq} = \begin{cases} \dfrac{\sum \nu_t^B}{\frac{1}{q} \sum_{j=0}^{q} \nu_{t-j}^B} & \text{if } \nu_t^B > \frac{1}{q} \sum_{j=0}^{q} \nu_{t-j}^S \\ \dfrac{\sum \nu_t^S}{\frac{1}{q} \sum_{j=0}^{q} \nu_{t-j}^S} & \text{if } \nu_t^S > \frac{1}{q} \sum_{j=0}^{q} \nu_{t-j}^B, \end{cases}$$

where q represents the length of the period used to compute the average trade size.

The volatility σ_{tq} is the standard deviation of the price log return at t over the last q intervals, where the return at t is

$$\rho_t = \log \left(\frac{\overline{P}_{t+1}^m}{\overline{P}_t^m} \right).$$

Finally, the variable δ_t indicates the sign of the change in mean of mid-price from time interval t to $t+1$:

$$\delta_t = \begin{cases} 1 \text{ if } \Delta_t > 0 \\ 0 \text{ otherwise.} \end{cases}$$

3.3 Model

We use a regression model to predict the impact of trading volume on the market price as described by Equation (1), in which Δ_t is the dependent variable, the features described in Table 2 are used as regressors, and ε_t is the residual error:

$$\begin{aligned} \Delta_t = \beta_1 V_t^O + \beta_2 P_t^O + \beta_3 V_t^T + \beta_4 D_{tq} \\ + \beta_5 \Delta_{t-1} + \beta_6 r_{tq} + \beta_7 \sigma_{tq} + \varepsilon_t \end{aligned} \tag{1}$$

Note that our model includes three volume variables, that is,

V_t^O, which is not widely used in the literature and is the average difference between the bid and ask side volumes of the limit order book during each time interval

V_t^T, which represents the net value difference between buy initiated trades and sell initiated trades fulfilled during an interval, and

r_{tq}, which compares the volume of trades during interval t to the average volume during the preceding period in the same direction.

Our model also includes two measures of volatility, that is,

σ_t, the standard deviation of the mid-quote returns

D_{tq}, the ratio of the average mid-price with respect to its moving average.

Finally, the model includes two measures of price variability:

Δ_{t-1}, the lagged mid-price change from interval t to $t+1$

P_t^O, the mean of difference of all quotes (both bid side and ask side) from their mid-price.

A common methodology in the literature to improve the forecasting precision of the price-impact model is to break the data set into two samples, yielding two regression models. The intuition is that price change in a bullish/bearish rally, or when there is sell/buy imbalance, could behave differently and it is plausible that impact of features on price change in these different market conditions could vary. Accordingly, we divide our data set into two parts, corresponding to either a positive or a negative price change (that is, according to the indicator variable δ_t. Tables 3 to 5 report the resuls of the regression for a 5-minute window (lenght of the t time intervals) and 10 look-back periods ($q = 10$); Table 3 displays the regression results for the complete data set, while Tables 4 and 5 report the results for positive and negative price changes, respectively.

Table 3: OLS regression results for 5-minute windows and 10 look-back periods (full data set).

Dep. Variable:	Δ	R-squared (uncentered):	0.106
Model:	OLS	Adj. R-squared (uncentered):	0.106
Method:	Least Squares	F-statistic:	293.4
Date:	Sat, 25 Dec 2021	Prob (F-statistic):	0.00
Time:	18:35:31	Log-Likelihood:	75008.
No. Observations:	17338	AIC:	-1.500e+05
Df Residuals:	17331	BIC:	-1.499e+05
Df Model:	7		

	coef	std err	t	P> |t|	[0.025	0.975]
β_1	0.0011	7.15e-05	15.004	0.000	0.001	0.001
β_2	6.9424	0.704	9.857	0.000	5.562	8.323
β_3	1.723e-05	6.28e-07	27.418	0.000	1.6e-05	1.85e-05
β_4	-0.0006	6.36e-05	-8.878	0.000	-0.001	-0.000
β_5	0.0989	0.009	11.562	0.000	0.082	0.116
β_6	0.0020	0.000	5.466	0.000	0.001	0.003
β_7	0.1349	0.013	10.176	0.000	0.109	0.161

Comparing the regression results in the three tables lead to the following observations:

1. The explanatory power of the model increases significantly when considering the direction in the price change, going from a R^2 value of 10.6% for the global

Table 4: OLS regression results for **positive** price changes over 5-minute windows and 10 look-back periods. Coefficients β_2 and β_5 do not appear since the corresponding variables were not significant and where removed from the regression.

Dep. Variable:	Δ	R-squared (uncentered):	0.581
Model:	OLS	Adj. R-squared (uncentered):	0.580
Method:	Least Squares	F-statistic:	2437.
Date:	Sat, 25 Dec 2021	Prob (F-statistic):	0.00
Time:	19:45:08	Log-Likelihood:	41531.
No. Observations:	8812	AIC:	-8.305e+04
Df Residuals:	8807	BIC:	-8.302e+04
Df Model:	5		

	coef	std err	t	P> |t|	[0.025	0.975]
β_1	0.0004	6.46e-05	5.992	0.000	0.000	0.001
β_3	5.685e-06	5.04e-07	11.280	0.000	4.7e-06	6.67e-06
β_4	-0.0003	5.93e-05	-5.671	0.000	-0.000	-0.000
β_6	0.0076	0.000	21.287	0.000	0.007	0.008
β_7	0.5644	0.012	48.019	0.000	0.541	0.587

Table 5: OLS regression results for **negative** price changes over 5-minute windows and 10 look-back periods.

Dep. Variable:	Δ	R-squared (uncentered):	0.567
Model:	OLS	Adj. R-squared (uncentered):	0.567
Method:	Least Squares	F-statistic:	1597.
Date:	Sat, 25 Dec 2021	Prob (F-statistic):	0.00
Time:	19:45:08	Log-Likelihood:	39901.
No. Observations:	8526	AIC:	-7.979e+04
Df Residuals:	8519	BIC:	-7.974e+04
Df Model:	7		

	coef	std err	t	P> |t|	[0.025	0.975]
β_1	0.0005	7.55e-05	6.758	0.000	0.000	0.001
β_2	6.7385	0.714	9.440	0.000	5.339	8.138
β_3	1.122e-05	6.69e-07	16.782	0.000	9.91e-06	1.25e-05
β_4	-0.0002	6.76e-05	-2.476	0.013	-0.000	-3.49e-05
β_5	0.0589	0.009	6.405	0.000	0.041	0.077
β_6	-0.0043	0.000	-10.966	0.000	-0.005	-0.004
β_7	-0.5070	0.014	-35.157	0.000	-0.535	-0.479

data set to values of 58.1% and 56.7% for positive and negative price changes, respectively.

2. Statistically significant features are not necessarily the same to explain positive and negative price changes.

3. While signs are the same, estimated coefficients can be significantly different according to the data set.

These results confirm the assumption of differing sensitivities to independent features according to market conditions, more specifically according to the direction of the price change from one time interval to the next[1].

3.4 Predicting price change directions

To complete the price-impact model, we need to be able to predict price change direction in order to be able to choose between the two sets of estimated coefficients. This issue can be related to a classification problem evaluating the probability of various possible outcomes for an event. In our case, the event is either an increase or a decrease in the mid-price from a given price interval to the next. Various methods exist to address classification problems, including logistic regression, support vector machines (SVM), decision trees, and random forests.

Table 6 compare four performance metrics obtained from these methods applied to a training subset of our data set, where

- Accuracy is the percentage of correctly predicted observations
- Precision is the ratio of correctly predicted positive observations to the total predicted positive observations
- Recall is the ratio of correctly predicted positive observations to all the positive observations
- WA is the weighted average of Precision and Recall.

[1] These observations are robust to the size of the time intervals and the number of time windows. The results presented here correspond to the setting yielding the best explanatory power for this data set.

Table 6: Classification perfomance metrics on the training data.

Method	Accuracy	Precision	Recall	WA
Logistic regression	61.61%	60.64%	65.80%	63.12%
SVM	62.23%	61.9%	63.3%	62.59%
Decision tree	59.07%	58.75%	60.48%	59.6%
Random forest	60.32%	60.6%	58.64%	59.6%

As these results suggest, all four methods perform comparably on the training data set, with logistic regression and SVM performing slightly better. Accordingly, we use logistic regression, which has the best Recall metric, for the production of price-change forecasts, which will be used for the investigation of spoofing activity in the next section. For a given given time interval, the price-change forecast is obtained in the following manner: if the logistic regression predicts an increase in price, then the forecast is obtained using the coefficients estimated from the "positive" regression model (Table 4); otherwise, the forecast is obtained using the coefficients of the "negative" regression model (Table 5).

4 Detecting spoofing

In this section, we first show how potential spoofing activities can be detected ex-post, using an illustrative example from a sample data point from the Coinbase exchange data set. We then discuss the possibility of using such a model to detect or prevent spoofing ex-ante.

4.1 Detecting spoofing ex-post

We start by defining two imbalance variables at a given time interval t. The "Order imbalance" variable is defined by

$$\gamma_t = \frac{\sum_{\tau=1}^{n_t} \sum_{i=1}^{m_\tau} \nu^b_{\tau i} - \nu^a_{\tau i}}{\sum_{\tau=1}^{n_t} \sum_{i=1}^{m_\tau} \nu^b_{\tau i} + \nu^a_{\tau i}} \qquad \in (-1, 1), \tag{2}$$

and the "Trade imbalance" variable is defined by

$$\lambda_t = \frac{\sum_{\tau=1}^{n_t} \nu_\tau^B - \nu_\tau^S}{\sum_{\tau=1}^{n_t} \nu_\tau^B + \nu_\tau^S} \qquad \in (-1, 1), \tag{3}$$

at time interval t.

The intuition behind these imbalance variables stems from the assumption of a healthy (non-spoofing) order book. If the order imbalance variable is high and close to 1 (*resp. low and close to -1*), we expect the trade imbalance for the very next intervals to increase (*resp. decrease*), since the traders are more interested in buying (*resp. selling*), and we expect the mid-price to increase (*resp. decrease*). If the reverse happens, then one may suspect a spoofing activity since the trade activities in the market are against the direction of the order book and price changes.

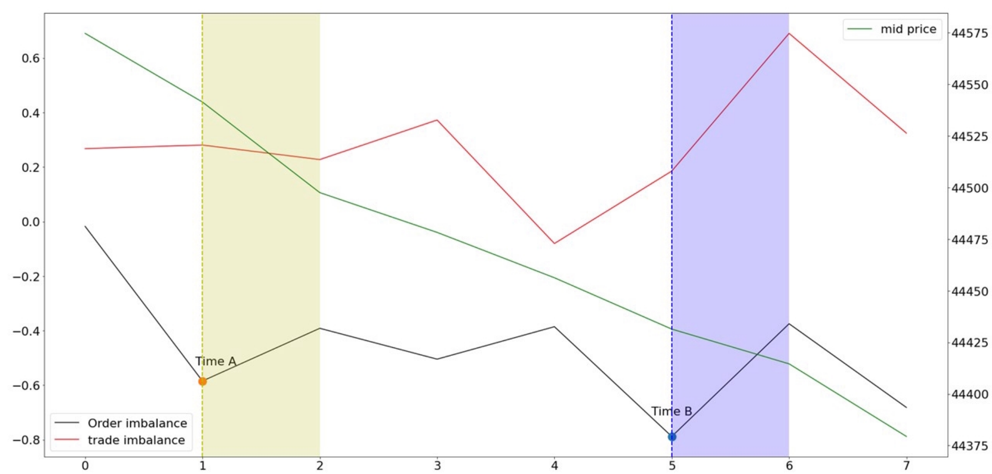

Figure 1: Illustration of ex-post spoofing detection.

We use Figure 1 to illustrate this intuition. At time labeled "A" in Figure 1, the order imbalance is negative, so that traders are more willing to sell and we expect the price and trade imbalance to decrease, which does happen in the time interval $[1, 2]$ (shaded in yellow). On the other hand, at time "B" where the order imbalance is very low, suggesting that the bid side volume of the order book is much stronger and traders are willing to sell, the mid-price decreases as expected, whereas the trade imbalance jumps during the time interval $[5, 6]$ (shaded in blue), indicating an increase in buy-initiated trades during the interval, against the expectation. As a result, one can suspect interval $[5, 6]$ to host spoofing activities.

Given this unexpected behavior, the trades and quotes during the interval following the suspicious point should be monitored. If a sudden cancellation in the order book and large trades fulfilled in the opposite direction are observed, then this suggests that there is an intention to inflate one side of the order book to manipulate the price and trade in the opposite direction, and trading accounts causing this chain of events are potential spoofers.

An illustrative example using a sample point in our data set is provided in Figure 2. This figure displays the imbalance variables and Bitcoin mid-price during 1-minute windows on February 28 between 14:40 and 15:15.

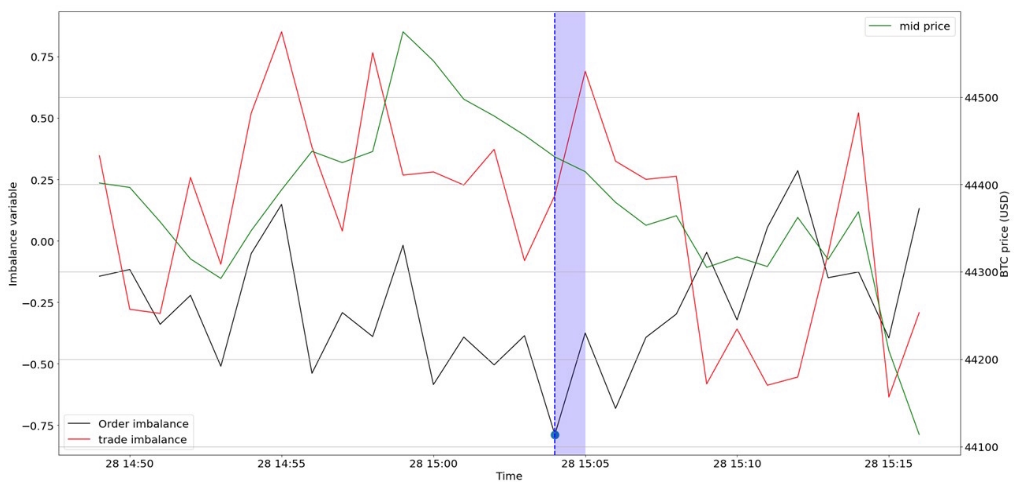

Figure 2: Illustration of possible spoofing activity in the Bitcoin market on February 28, 2021.

According to the value of the order and trade imbalance variables during this period, the shaded window starting at 15:04 in Figure 2 is suspicious for spoofing. Accordingly, we investigate the trades in the next one-minute interval. Note that we use short time intervals and windows is that spoofing is usually implemented using algorithmic trading agents interacting with other algorithmic strategies that account for the smallest updates in the order book.

Figure 3 displays the volume distribution of buy-initiated trades fulfilled during the next one-minute interval (15:04 to 15:05). We only consider buy-initiated trades since the order imbalance is very low while the trade imbalance is increasing, indicating that any potential spoofer would be buying the asset.

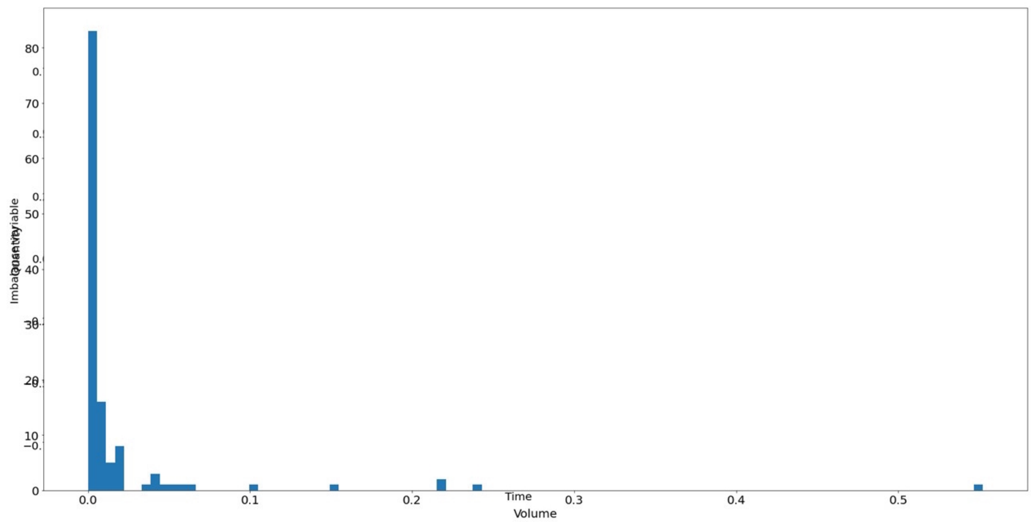

Figure 3: Volume distribution of buy-initiated trades in the time-window from 15:04 to 15:05.

As Figure 3 suggests, there is an outlier trade, on which we are going to focus. Table 7 represents the detailed attributes of this trade, which shows that it has been fulfilled at 15:04:08. As a result, we are going to examine the order book data just before this instant.

Table 7: Detail attributes of outlier trade

Attribute	Value
time	2021-02-28 15:04:08
size	0.552009
price	44442.8
direction	buy

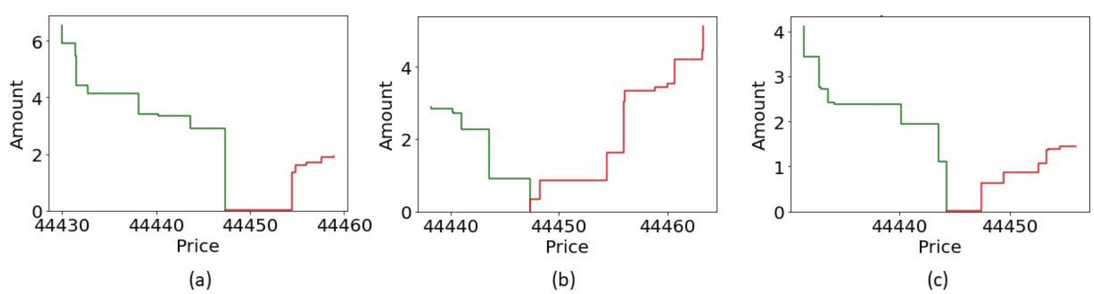

Figure 4: Order book visualization, where bid quotes are displayed in green and ask quotes are displayed in red. Panel (a): 15:03:43. Panel (b): 15:03:43. Panel (c): 15:04:03.

Figure 4 displays the updates in the order book prior to the outlier trade. The green lines display bid quotes and red lines display ask quotes in the order book. At 15:03:43, the order imbalance is toward bid quotes, whereas at 15:03:53 the volume of ask quotes increases and the order imbalance decreases to a negative value, showing more supply in the order book, which results in the fall in mid price at 15:04:03. This situation could be a good justification for spoofing: at 15:03:43 the best ask price is high and the spoofer is looking for a lower price, so the volume on the ask side is inflated to decrease the price; as soon as the price reaches the desired level (at 15:04:03), a buy trade is fulfilled by the spoofer. While it is possible that this particular trade be due to spoofing (or layering), there is no absolute certainty. Specifically:

- The granularity of the order book data is vital, since most spoofing activities happen in very short periods, so that instantaneous monitoring is required from regulators.

- It is necessary to know the account information of traders, that is, which trader has posted which quotes. In the case of a cancellation, subsequent activities by the same trader could then be monitored by financial authorities.

4.2 Detecting spoofing ex-ante

The previous sections have shown that market information, including quotes price and volume, has significant impact on price change. The following observations can be used to avoid spoofing ex-ante using information from the market-impact model:

- Since spoofers manipulate and trade in short periods, price-impact regression models should be performed over short time windows.
- The regulators should set a limit for price changes during a given time interval.
- All regression variables except for quotes' volume should be collected from the market information.
- Every coming order should be checked using the regression models; if the predicted price change is beyond the defined limit, that quote should be declined or activities by that trading account should be monitored.

5 CONCLUSIONS

In this paper, we propose two approaches for detecting spoofers in financial markets. The ex-post method was tested on a sample trade that could be a candidate for spoofing. However, it is not possible to investigate this trade any further since the data for trading accounts and their activities were not available. An ex-ante approach was also proposed to detect spoofers prior to an illegal activity based on their predicted impact on market price measured by the results of a 2-step regression.

ACKNOWLEDGEMENTS

This research was supported by the Centre d'intelligence en surveillance des marchés financiers de l'ESG UQAM. Funding and student support from the CISMF is gratefully aknowledged.

References

[1] Allen, Franklin, and Douglas Gale. Stock-price manipulation. *The Review of Financial Studies* 5.3 (1992): 503-529.

[2] Aloosh, Arash, and Jiasun Li. Direct evidence of bitcoin wash trading. *SSRN 3362153* (2022).

[3] Cao, Yi, Yuhua Li, Sonya Coleman, Ammar Belatreche, and Thomas Martin McGinnity. Detecting price manipulation in the financial market. In *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics* (CIFEr) (2014): 77-84.

[4] Cartea, Álvaro, Sebastian Jaimungal, and Yixuan Wang. Spoofing and price manipulation in order-driven markets. *Applied Mathematical Finance* 27.1-2 (2020): 67-98.

[5] Chen, Weili, YueJin Xu, Zibin Zheng, Yuren Zhou, Jianxun Eileen Yang, and Jing Bian. Detecting "Pump & Dump" Schemes on cryptocurrency market using an improved Apriori Algorithm. *IEEE international conference on service-oriented system engineering* (SOSE) (2019): 293-2935.

[6] Cong, Lin William, Xi Li, Ke Tang, and Yang Yang. Crypto wash trading. *Management Science* 69, no. 11 (2023): 6427-6454.

[7] Eigelshoven, Felix, Andre Ullrich, and Douglas A. Parry. Cryptocurrency market manipulation: A systematic literature review. *International Conference on Information Systems*. 2021.

[8] Gandal, Neil, J. T. Hamrick, Tyler Moore, and Tali Oberman. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics* 95 (2018): 86-96.

[9] Griffin, John M., and Amin Shams. Is Bitcoin really untethered? *The Journal of Finance* 75.4 (2020): 1913-1964.

[10] Hamrick, J. T., Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. An examination of the cryptocurrency pump-and-dump ecosystem. *Information Processing & Management* 58-4 (2021): 102506.

[11] Lee, Eun Jung, Kyong Shik Eom, and Kyung Suh Park. Microstructure-based manipulation: Strategic behavior and performance of spoofing traders. *Journal of Financial Markets* 16.2 (2013): 227-252.

[12] Le Pennec, Guénolé, Ingo Fiedler, and Lennart Ante. Wash trading at cryptocurrency exchanges. *Finance Research Letters* 43 (2021): 101982.

[13] Mansourifar, Hadi, Lin Chen, and Weidong Shi. Hybrid cryptocurrency pump and dump detection. arXiv preprint arXiv:2003.06551 (2020).

[14] Twomey, David, and Andrew Mann. Fraud and manipulation within cryptocurrency markets. *Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation* 624 (2020).

[15] Xu, Jiahua, and Benjamin Livshits. The anatomy of a cryptocurrency {Pump-and-Dump} scheme. *28th USENIX Security Symposium* (2019).

[16] Wang, Y. Strategic spoofing order trading by different types of investors in the futures markets. *Wall Street Journal* (2015).