Digital Finance (2023) 5:519–562
https://doi.org/10.1007/s42521-023-00095-9

ORIGINAL ARTICLE

Market impact and efficiency in cryptoassets markets

Emilio Barucci[1] · **Giancarlo Giuffra Moncayo**[1] · **Daniele Marazzina**[1]

Received: 23 June 2022 / Accepted: 28 August 2023 / Published online: 16 September 2023
© The Author(s) 2023

Abstract
We analyze markets for cryptoassets (cryptocurrencies and stablecoins), investigating market impact and efficiency through the lens of the market order flow. We provide evidence that markets where cryptoassets are exchanged between themselves play a central role on price formation and are more efficient than markets where cryptocurrencies are exchanged with the US dollar. For the first set of markets we observe some evidence of the presence of insiders/contrarians, instead in the latter we observe the predominance of herding and trend-followers.

Keywords Stablecoin · Cryptocurrency · Liquidity · Efficiency · Market impact

JEL Classification E41 · E50

1 Introduction

Cryptoassets represent an interesting laboratory setting for financial market research. They are traded in anonymous exchanges, several exchanges are available for each market, there are no best price execution obligations, few traders actively operate in several exchanges, price discovery and information flow are not smooth as in regulated markets. These features render the analysis of efficiency and deepness of these markets a very interesting—and largely unexplored—topic.

Differently from many papers that concentrate only on markets where cryptoassets are exchanged with the US dollar, see for example Baur et al. (2018), Brandvold et al. (2015), Grobys (2021), Lintilhac & Tourin (2017), Petukhina et al. (2021), we consider markets where fiat money is exchanged with cryptoassets as well as

✉ Daniele Marazzina
daniele.marazzina@polimi.it

Emilio Barucci
emilio.barucci@polimi.it

Giancarlo Giuffra Moncayo
giancarlo.giuffra@polimi.it

[1] Department of Mathematics, Politecnico di Milano, 32 p.zza L. da Vinci, Milan, Italy

markets where cryptoassets are exchanged with each other (including Tether). We would like to stress the importance of considering the latter markets: their role is becoming more and more relevant, for example Ciaian et al. (2018) show that 47% of Ethers are exchanged with Bitcoin, and only 37% with US dollar or Euro.

We investigate market impact and market efficiency (return and arbitrage opportunities) at high frequency, centering the analysis on market order flow (signed volume). We show that markets where cryptoassets are exchanged between themselves play a central role on price formation and are more efficient than markets where cryptocurrencies are exchanged with the US dollar, where there is a predominance of herding.

More in details, analysing the order flow, we show that traders act as trend-followers at high frequency in all markets and behave as contrarians at daily frequency (as in stock exchange) only in markets where cryptoassets are exchanged between themselves. There is some evidence that sophisticated traders operate in these markets exploiting financial time series regularities for trading, whereas in the second set of markets traders mostly herd at high frequency with low persistency of order flow at lower frequencies.

Market impact refers to the effect of the order flow on the contemporaneous price change/return. The topic has been addressed in a large literature for stock exchange and exchange rate markets, see Berger et al. (2008), Chordia et al. (2002, 2005), Cont et al. (2013), Evans & Lyons (2002), and in cryptocurrency markets, see Donier & Bonart (2015), Lyons & Viswanath-Natraj (2019), Makarov & Schoar (2020), Silantyev (2019). In agreement with the stock exchange literature, we observe that there is a positive relation between the order flow and the contemporaneous asset return in all markets, however the market impact of the order flow is negligible for markets where cryptocurrencies are exchanged with the US dollar and high for markets where cryptocurrencies are exchanged with Tether. Moreover, in the latter markets there is some evidence that sophisticated traders operate with an inventory target, whereas the order flow does not seem to contain relevant information in markets where cryptoassets are exchanged with the US dollar.

As far as market efficiency is concerned, we investigate the topic along two different directions: return predictability and arbitrage opportunities.

Market efficiency of cryptoassets has been investigated in many papers through statistical tests, see, e.g., Bariviera (2017); Brauneis and Mestel (2018); Nadarajah and Chu (2017); Sensoy (2019); Tiwari et al. (2018); Urquhart (2016). In our analysis we take a different perspective. Given that the order flow has a market impact on contemporaneous price movements/returns, we follow Chordia et al. (2002) and Chordia et al. (2005) testing for the random walk hypothesis at different frequencies, i.e., return does not allow to predict future return, adding the lagged order flow as explanatory variable. In the stock market, the authors find that the order flow has an impact on future market return over a short horizon, then sophisticated traders react to order imbalances within trading day by undertaking countervailing trades exploiting serial correlation of the order flow. As a result, the order flow affects market returns over short horizons (5 min–1 h) but not over the day. We find evidence that markets where cryptocurrencies are exchanged with the US dollar are strongly inefficient, the others being more efficient.

We build on the fact that cryptoassets are traded in several exchanges to investigate the possibility that arbitrage opportunities arise by trading on different exchanges (the same pair of cryptoassets) or trading different pairs of cryptoassets.[1] The issue has been already investigated by Makarov and Schoar (2020) analyzing the existence of arbitrage opportunities in the Bitcoin–US dollar market. They show that there are large deviations in Bitcoin prices across exchanges that often persist for a long time. However, arbitrage opportunities are limited unless different currencies are involved. We show that in the very short term trading activity does not close an arbitrage opportunity, at a high frequency there is a continuation/amplification effect. Then traders discover arbitrage opportunities and the market moves in the direction of closing arbitrage opportunities. Crypto-markets are homogenous on this point. However, arbitrage opportunities are closed more quickly in markets where Tether is involved rather than in markets where cryptocurrencies are exchanged with the US dollar.

These results suggest that markets where cryptoassets are exchanged with each other play a central role on price formation. In these markets there are sophisticated traders who ease the aggregation of opinions/technology shocks. Instead, there is a predominance of herding in markets where cryptoassets are exchanged with the US dollar. The result is confirmed by observing that the order flow in the latter markets does not seem to contain relevant pieces of information and that they are strongly inefficient, whereas the other markets are more efficient.

The paper is organized as follows. In Sect. 2 we describe the dataset of our analysis. In Sect. 3 we provide a statistical analysis of the order flow time series. In Sect. 4 we investigate market impact of order flow. In Sect. 5 we investigate efficiency of crypto-markets. In Sect. 6 we analyze the profitability of arbitrage strategies. In Appendix 1 we provide the list of exchanges considered in the analysis for each pair of cryptoassets. In Appendix 1 we describe in detail how arbitrage opportunities are constructed. In Appendix 1 we provide a table showing the autocorrelation of time series.

2 The dataset

We start defining the main quantities considered in this work: dealing with the European Central Bank's definition,[2] a cryptoasset is an asset recorded in digital form and enabled by the use of cryptography that is not and does not represent a financial claim on, or a liability of, any identifiable entity. In the following, we define a cryptocurrency as a native blockchain cryptoasset. As an example, Ether is a cryptocurrency since it is the native digital asset of Ethereum (fees to use Ethereum must be paid in Ether), while Tether, a stablecoin, is a cryptoasset, but not a cryptocurrency, since it is a digital token exchanged on several blockchains, like Ethereum or Tron, none of them using Tether as the native digital asset. In the following we also define

[1] The presence of arbitrage opportunities has been investigated in stock markets considering dual listed stocks, see (Ghadhab & Hellara, 2015; De Jong et al., 2009) and in ETF values, see (Marshall et al., 2013).

[2] https://www.ecb.europa.eu/paym/intro/mip-online/2019/html/1906_crypto_assets.en.html.

Fig. 1 Bitcoin value in USD on the period April 1, 2019–October 31, 2020

a crypto-market as a market where a pair of cryptoassets, or one cryptoasset and a currency, are traded.

We focus our analysis on the period April 1, 2019–October 31, 2020, see Fig. 1 for the US dollar price of Bitcoin during the period of analysis. The main reason for considering this period is that it is just before the huge surge in 2021: as crypto-markets are stable we may investigate their functioning concentrating on market impact and efficiency abstracting from speculative and herding phenomena that are likely to characterize crypto-markets afterwards.

We refer to a pair/market for the currency and asset involved, e.g., BTC-USD stands for the Bitcoin–US dollar pair. The BTC–USD price represents the amount of USD necessary to buy/sell a BTC. Each pair of currency/assets is traded in several exchanges. We consider markets involving currencies, cryptocurrencies, and stablecoins. As far as currency is concerned, we only deal with the US dollar (USD). We consider the two most relevant cryptocurrencies by market capitalization: Bitcoin (BTC) and Ether (ETH). We restrict our attention to the stablecoin with the largest market capitalization: Tether (USDT), a stablecoin pegged to the US dollar and collateralized by the US dollar itself (fiat-backed stablecoin). The choice of concentrating on these cryptoassets is motivated by their relevance. Bitcoin and Ether are the two cryptocurrencies with the largest capitalization in US dollar from February 2016 (few months after the launch of Ethereum) to nowadays. Trading volumes of markets where Tether is exchanged against major cryptocurrencies have steadily grown with Tether becoming the cryptoasset with the largest trading volume since the second quarter of 2019. Also its market capitalization increased significantly: during the period covered in our analysis, Tether capitalization went from 2 billion dollars to 16 billion. We only deal with markets where cryptoassets are traded with

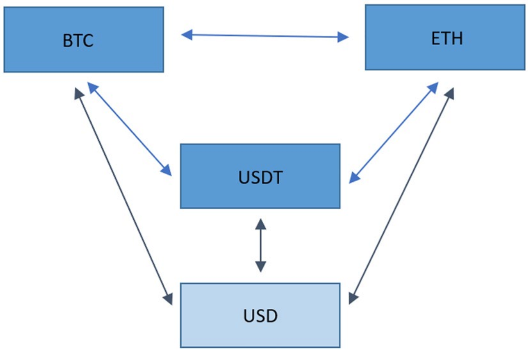


Fig. 2 Pairs considered in the analysis

the US dollar, the rationale of this choice is that these are the most liquid markets and therefore the most significant markets for our analysis. Moreover, considering different currencies, we should consider the exchange rate between the currencies complicating significantly the picture; for an analysis of arbitrage opportunities dealing with different currencies we refer to Makarov and Schoar (2020).

We consider markets where cryptoassets are exchanged between themselves, and markets where cryptoassets are exchanged with the US dollar. We end up with three different sets of pairs: BTC and ETH against USDT (two pairs), ETH against BTC (one pair), BTC, ETH, USDT against USD (three pairs), see Fig. 2. The six pairs are associated to twentyone exchanges, see Appendix 1 for the list of exchanges for each pair. Exchanges have been selected to cover at least 70% of each market according to coinmarketcap.com data regarding the trading volume as in September 2020.[3] In Sect. 3–5 we aggregate information from all the exchanges of each market, and therefore we look at all the exchanges as a unique market. In Sect. 6 we deal with data at exchange level.

Differently from other papers on cryptocurrencies we deal with cryptocurrencies and stablecoins markets. The reason is that stablecoins are becoming more and more important, playing a relevant role on trading cryptoassets, see also the discussion in Barucci et al. (2022). To better understand the nexus among standard currency, cryptocurrencies and stablecoins we consider the following example. A person holding BTC on a wallet, sells them against USD in Exchange A, and uses USD to acquire Ether (ETH) in Exchange B. To deploy these trades, the steps could be as follows: the person

1. sends BTC to Exchange A, and sells them against USD. Then asks Exchange A to transfer the acquired USD to a bank account;
2. transfers USD to Exchange B (via bank transfer or credit card);

[3] Only exchanges with a CoinMarketCap Confidence Indicator equal to High were considered. CoinMarketCap exploits a machine learning model to estimate volume of every single market pair that reports exchanges. Given the estimated volume, they detect an outlier if the exchange reports far higher volume than the model predicts, allowing to flag them accordingly to their Confidence indicator. A high confidence indicator corresponds to high level of confidence in the market's reported volume. https://coinmarketcap.com/.

3. acquires ETH against USD and asks Exchange B to move ETH to a wallet.

The bank transfer from Exchange A to the agent can take a significant time delay, and usually there are high fees. An alternative approach would be to leave a certain amount of USD deposited on different exchanges, to be used for trading, but this approach is inefficient as it requires a significant amount of USD to be allocated in the exchanges.

A shortcut is provided by stablecoins, tokens that have been introduced to capitalize the benefits of cryptocurrencies along with price stability. Exploiting stablecoins the transactions involved in the second example can be carried out as follows: the person

1. sends BTC to Exchange A, and sells them against a stablecoin. Then asks Exchange A to move the stablecoins to a wallet;
2. transfers stablecoins from the wallet to Exchange B;
3. acquires ETH against stablecoins and asks Exchange B to move ETH to the wallet.

In this case no bank transfer is necessary.

This example highlights the relevance of stablecoins in cryptoasset markets to facilitate transactions of cryptoassets without involving USD, i.e., when trades only occur in the cryptoassets domain. Because of these features, we claim that markets involving stablecoins play a relevant role for sophisticated traders who want to detain cryptocurrencies for technology or liquidity reasons.

The dataset is made up of tick-by-tick trading information obtained from Kaiko.[4] We emphasize that our dataset represents the registered trading activity occurring in the different exchanges with synchronous trading/price observations. The dataset captures actual trading activity and does not look at blockchain activity. On the importance of the right choice of the data provider and of the use of tick-by-tick data in cryptocurrency markets, we refer to Alexander and Dakos (2020); Manahov (2021); Vidal-Tomás (2021).

For each transaction, trade information includes the following items: Exchange, Currency/asset pair, Date (timestamp in milliseconds), Price of the transaction in the reference currency, Amount (quantity of the asset), Sell (True or False, referring to the trade direction, a trade marked as 'true' means that a price taker placed a market sell order).

We deal with outliers applying a variation of the methodology proposed in Brownlees and Gallo (2006). For each point-observation of the raw high frequency time series, we consider the interval of 60 s centered on that point and exclude it if the price is more than three standard deviations away from the mean of the interval. The observation is discarded for all time series. For all pairs except USDT-USD,

[4] Kaiko has been collecting trading information about cryptocurrencies since 2014, it provides data for more than 100 exchanges and more than 70 000 currency pairs. https://www.kaiko.com/.

Table 1 Statistics for 1 min log-returns

Symbol	Mean	Stddev	Skewness	Kurtosis	Min	Max
BTC-USDT	9.19×10^{-7}	0.0011	− 0.16	188.26	− 0.06	0.07
ETH-USDT	1.93×10^{-6}	0.0013	0.05	215.49	− 0.08	0.08
ETH-BTC	9.05×10^{-8}	0.0008	− 0.53	160.18	− 0.06	0.05
BTC-USD	-5.54×10^{-5}	0.0049	− 1.30	163.77	− 0.42	0.14
ETH-USD	-6.64×10^{-5}	0.0069	− 0.03	56.24	− 0.16	0.12
USDT-USD	-2.14×10^{-6}	0.0028	1.02	331.21	− 0.08	0.08

Number of samples: 833,760

less than 0.01% of the original sample was discarded. For USDT-USD the fraction was around 5%.

Prices are sampled at 1 s frequency. For each second interval, we compute the price as the average price of trades executed during that second, the prices are weighted by the volume of the corresponding trades. Starting from the price sampled at 1 s frequency, we compute the one minute log-return. For each minute t, where there is at least one executed trade, we identify the first second within that minute with a transaction. We denote the price of that transaction as p_t and the log-return r_t for minute t is computed as $r_t = \log\left(\frac{p_{t+1}}{p_t}\right)$. If no trade is executed during minute t, then $r_t = 0$.

In Table 1 we report some basic statistics on returns. Notice that the average log-return is positive when cryptoassets are exchanged with other cryptoassets and negative when they are exchanged with the USD. Skewness is limited with the exception of USDT-USD and BTC-USD markets. In the latter market we observe a negative value showing the relevance of abrupt negative returns. Kurtosis is high in all the markets. Notice that the standard deviation of returns in markets involving USD is much higher than in the three markets involving only cryptoassets.

3 Order flow

Our analysis is centered on the market Order Flow (OF) or market imbalance. In Chordia et al. (2005) three different specifications of OF are considered: the number of buyer-initiated less the number of seller-initiated trades, the number of buyer-initiated shares purchased less the number of seller-initiated shares sold, the dollars paid by buyer-initiators less the dollars received by seller-initiators. In what follows, we consider the OF in terms of signed volume (the second specification) as it is considered in other papers on cryptocurrencies, see Silantyev (2019), Lyons & Viswanath-Natraj (2019), while the third specification is considered in Makarov and Schoar (2020).[5]

[5] Silantyev (2019) compares the market impact of the OF computed from trading activity (trade flow imbalance) to the one proposed by Cont et al. (2013) on the order book imbalance (order flow imbalance), i.e., variation in the difference of the orders on the bid and ask side of the market (best bid and

Table 2 Statistics for OF at 1 min frequency

Symbol	Mean	Stddev	Skewness	Kurtosis	Min	Max
BTC-USDT	– 13829.53	63,6071.06	– 2.18	202.35	$-\ 5.43 \times 10^7$	2.56×10^7
ETH-USDT	– 7316.94	199,468.53	– 2.66	315.80	-1.33×10^7	1.10×10^7
ETH-BTC	728.89	116,323.78	– 0.99	94.39	-7.56×10^6	4.09×10^6
BTC-USD	8338.23	350,093.51	– 2.99	283.21	-2.48×10^7	1.86×10^7
ETH-USD	– 573.26	102,771.34	– 6.57	646.60	-1.06×10^7	6.08×10^6
USDT-USD	216.83	42,372.78	– 7.11	1574.08	-3.96×10^6	3.61×10^6

To facilitate the comparison, the order flow is evaluated in US dollars using the average daily rate. Number of samples: 833,760

The order flow is computed at the one minute frequency and it is defined as the buyer-initiated volume minus the seller-initiated volume:

$$OF = \sum_{i=1}^{n} V_i \cdot S_i,$$

where V_i is the volume of the i-th trade and S_i denotes the market side initiating the trade: 1 for the buyer, -1 for the seller. n denotes the number of trades in the minute. Some basic statistics of OF are reported in Table 2.

To investigate how OF affects market activity, in Table 3 we first provide results of the Ordinary Least Squares (OLS) regression of OF_t on lagged return (r_{t-1}), lagged order flow (OF_{t-1}), and on both lagged variables ($r_{t-1},\ OF_{t-1}$), that is

$$OF_t = \beta_0 + \beta_1 OF_{t-1} + \epsilon_t, \tag{1}$$

$$OF_t = \beta_0 + \beta_2 r_{t-1} + \epsilon_t, \tag{2}$$

$$OF_t = \beta_0 + \beta_1 OF_{t-1} + \beta_2 r_{t-1} + \epsilon_t, \tag{3}$$

ϵ being the error random variable.[6] In the analysis, we deal with the 1, 5 and 10 min, hour, and daily frequency. In Table 11 of Appendix 1 we also report autocorrelations at different frequencies (5 min, 1 h, and 1 day, for the sake of brevity we omit the 1 and 10 min frequency information).

Let us consider the first column for each frequency in Table 3, i.e., regression (1), as well as Table 11 on OF. The autoregressive component for the OF is positive for all the markets. The pattern with respect to the sampling frequency looks different: as it decreases, the autoregressive coefficient decreases for BTC-USD and ETH-USD markets and increases for BTC-USDT, ETH-USDT, ETH-BTC markets, being

Footnote 5 (continued)

best ask). He concludes that trade flow imbalance explains a larger fraction of price variations in the BTC-USD market.

[6] In all the regressions, all the variables are normalized and therefore the intercept is not reported.

Table 3 OLS regression of OF_t on lagged return (r_{t-1}), lagged order flow (OF_{t-1}), and both lagged variables ($r_{t-1},\ OF_{t-1}$) at 1, 5 and 10 min, hour, daily frequency

Explanatory variable	OrderFlow														
	1 min			5 min			10 min			1 h			1 day		
	BTC-USD														
Order-Flow$_{t-1}$	0.2330		0.2313	0.1960		0.1950	0.1824		0.1808	0.1393		0.1404	0.0076		0.0077
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(0.8561)		(0.8551)
Log-return$_{t-1}$		0.0429	0.0134		0.0457	0.0085		0.0512	0.0156		0.0117	− 0.0136		− 0.0211	− 0.0211
		(< 0.001)	(< 0.001)		(< 0.001)	(0.0071)		(< 0.001)	(0.0011)		(0.3408)	(0.2687)		(0.5363)	(0.5364)
R^2	5.4299	0.1838	5.4476	3.8410	0.1250	3.8452	3.3285	0.1386	3.3410	1.9392	0.0066	1.9480	0.0058	0.0674	0.0733
	ETH-USD														
Order-Flow$_{t-1}$	0.2108		0.2103	0.1907		0.1901	0.1941		0.1932	0.1442		0.1439	0.0751		0.0761
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(0.0733)		(0.0699)
Log-return$_{t-1}$		0.0247	0.0063		0.0309	0.0071		0.0380	0.0126		0.0240	0.0049		− 0.0252	− 0.0272
		(< 0.001)	(< 0.001)		(< 0.001)	(0.0224)		(< 0.001)	(0.0065)		(0.0500)	(0.6841)		(0.4817)	(0.4472)
R^2	4.4430	0.0608	4.4468	3.6364	0.0590	3.6394	3.7681	0.0799	3.7768	2.0791	0.0281	2.0803	0.5636	0.0872	0.6649
	BTC-USDT														
Order-Flow$_{t-1}$	0.1724		0.1550	0.1959		0.1828	0.2170		0.2001	0.2044		0.1902	0.3312		0.5082
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
Log-return$_{t-1}$		0.1088	0.0386		0.1416	0.0321		0.1821	0.0447		0.2286	0.0451		0.0088	− 0.7795
		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(0.0103)		(0.9359)	(< 0.001)
R^2	2.9737	1.1844	3.0921	3.8381	1.3673	3.8913	4.7100	1.8285	4.7916	4.1796	1.7070	4.2257	10.9573	0.0011	16.7106
	ETH-USDT														

Table 3 (continued)

Explanatory variable	OrderFlow														
	1 min			5 min			10 min			1 h			1 day		
Order-Flow$_{t-1}$	0.1792		0.1785	0.2111		0.2044	0.2566		0.2407	0.2561		0.2436	0.4416		0.6531
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
Log-return$_{t-1}$		0.0902	0.0015		0.1467	0.0159		0.2134	0.0415		0.3062	0.0413		− 0.0582	− 1.2650
		(< 0.001)	0.2213		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(0.0314)		(0.6778)	(< 0.001)
R^2	3.2120	0.8130	3.2122	4.4493	1.4161	4.4615	6.5760	2.4506	6.6431	6.5553	2.6209	6.5871	19.3336	0.0304	29.2566
	ETH-BTC														
Order-Flow$_{t-1}$	0.0947		0.0890	0.1289		0.1292	0.1545		0.1618	0.3312		0.3493	0.8615		0.8654
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
Log-return$_{t-1}$		0.0453	0.0203		0.0417	− 0.0014		0.0278	− 0.0298		0.0196	− 0.1171		− 0.0416	− 0.2758
		(< 0.001)	(< 0.001)		(< 0.001)	(0.6332)		(< 0.001)	(< 0.001)		(0.1493)	(< 0.001)		(0.8204)	(0.0030)
R^2	0.8974	0.2056	0.9354	1.6608	0.1358	1.6609	2.3879	0.0531	2.4437	10.9717	0.0152	11.4801	74.1981	0.0091	74.5951
	USDT-USD														
Order-Flow$_{t-1}$	0.3182		0.3182	0.3451		0.3451	0.2554		0.2554	0.2327		0.2326	0.3330		0.3331
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
Log-return$_{t-1}$		− 0.0030	− 0.0024		− 0.0066	− 0.0063		− 0.0011	− 0.0019		0.0199	0.0161		0.0067	0.0221
		(0.1361)	(0.1988)		(0.1008)	(0.0940)		(0.8736)	(0.7651)		(0.2635)	(0.3508)		(0.9600)	(0.8600)
R^2	10.3799	0.0007	10.3804	12.1094	0.0023	12.1115	6.6453	(< 0.001)	6.6454	5.4249	0.0095	5.4312	11.0206	0.0004	11.0255

Each cell reports the regressor coefficient and the corresponding p-value (below the coefficient, in brackets). R^2 is shown in percentage points

almost constant for USDT-USD. At the daily frequency it is limited and weakly or non statistically significant for BTC-USD and ETH-USD markets. The coefficient is high, positive and statistically significant at any frequency for BTC-USDT, ETH-USDT, BTC-ETH, and USDT-USD markets. The explanatory power of the regression (R^2) decreases and becomes negligible as frequency decreases for the first set of markets and increases for the latter. Results are confirmed controlling for lagged return in the regression (3) (third column).

The relation between OF and lagged return, i.e., regressions (2)–(3), depends on the frequency and on the market. Notice that a negative coefficient would suggest that traders are contrarians either for liquidity or information/arbitrage arguments, instead a positive coefficient would suggest that traders are trend-followers either for the spread of information or because they are herding. We concentrate on the regression (3) (third column for each frequency) which yields a higher R^2 with respect to the regression only involving r_{t-1}. In the USDT-USD market we observe a non statistically significant coefficient for lagged return at any frequency. The result is likely to be due to the features of USDT: being pegged to one USD, USDT is characterized by a constant fundamental value with no dissemination of new information about it and therefore there is no economic rationale for traders to act as contrarians/trend-followers reacting to past returns. On the possibility of trading motivated by arbitrage arguments (with respect to the conversion value) see Sect. 6. In BTC-USD and ETH-USD markets we observe a positive coefficient for r_{t-1} up to the ten minute frequency, then the coefficient of the linear regression turns out to be non statistically significant and also the explanatory power of the regression becomes negligible.

Considering markets where cryptoassets are exchanged with each other, we observe a positive coefficient for r_{t-1} at high frequency up to 1 h (BTC-USDT and ETH-USDT) and one minute (ETH-BTC), then the coefficient becomes negative; coefficients are statistically significant with only few exceptions.

Results look different from what is obtained for stock exchanges. Chordia et al. (2002) showed that the order flow in stock exchanges is highly persistent at daily frequency and that investors in aggregate are contrarians: they buy after market declines and sell after market moves up. Chordia and Subrahmanyam (2004) provided a theoretical model replicating the above regularities considering informed and discretionary liquidity traders (they can split their order in different periods). Notice that in a cryptoasset market we cannot assume that there are informed traders, as a matter of fact it is difficult to define its fundamental value as there is no cash flow associated with a cryptocurrency (fiat money). The exception is provided by the USDT-USD market: being USDT pegged to one US dollar, its fundamental value is well defined and does not vary over time. In the other markets we may only assume that there are insiders and outsiders with different opinions on the technology.

The analysis of crypto-markets provides different results. BTC-USD and ETH-USD markets are characterized by very short effects. The order flow is characterized by a strong positive autoregressive component over a short time window (up to 1 h in our analysis) coupled with a positive effect associated with past return; over a longer time window (1 day) both the the autoregressive component of the order flow and past returns are not statistically significant. These results highlight that the

Fig. 3 Scatter plot of BTC-USD (left) and BTC-USDT (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: OF, *y*-axis: log-return. The straight line obtained from the linear regression is reported ▶

markets are characterized by herding effects at high frequency, confirming the analysis in Ballis and Drakos (2020); Bouri et al. (2019); King and Koutmos (2021); Manahov (2021), with no evidence of countervailing-contrarian forces over the day as observed in stock markets. BTC-USDT, ETH-USDT and ETH-BTC markets look different: serial correlation of OF is positive, statistically significant and increases as the frequency decreases, the relation with past return shows that traders act as trend-followers at high frequency and as contrarians at low frequency.

We interpret this evidence as showing that sophisticated traders looking at exploiting financial time series regularities for trading are present only in markets where cryptoassets are exchanged between themselves and not in markets where cryptoassets are exchanged with the USD. In the latter markets, traders mostly herd at high frequency with low persistency of order flow at lower frequencies.

These results confirm the analysis provided by Barucci et al. (2022) showing that markets where a cryptoasset is exchanged against the US dollar play a less signifcant role with respect to markets where cryptoassets are exchanged between themselves. The second set of markets seems to be the place where prices are formed aggregating preference/technology shocks and heterogeneous opinions. In particular, the BTC-USDT market represents a privileged locus for price aggregation and not only for manipulation of BTC as shown in Griffin and Shams (2020).

4 Market impact

We investigate price pressure in crypto-markets, i.e., the effect of the order flow on market return. The literature on stock exchanges has shown that the order flow affects the contemporaneous market return, see Chordia et al. (2002, 2005), Cont et al. (2013). In Table 4 we provide results on a regression of the log-return at time t (r_t) on OF_t, that is

$$r_t = \beta_0 + \beta_1 OF_t + \epsilon_t,$$

at 1, 5 and 10 min, as well as 1 h and 1 day frequency. Results are also reported in Fig. 3, where we plot log-return against OF together with the line obtained from the linear regression for two representative markets (BTC-USD and BTC-USDT) at 1, 5 and 10 min, 1 h, and 1 frequency, see Appendix 1 for the other markets.

We observe a positive statistically significant effect in all the markets with the exception of USDT-USD (all frequencies) and of ETH-BTC, BTC-USD, ETH-USD at the daily frequency (positive but not significative).[7] The results are confirmed looking at the coefficient of OF_t in all regressions reported in Table 4 including OF_{t-1}, that is

[7] Results are also graphically illustrated in Fig. 3 and in Appendix 1, Figs. 6, 7.

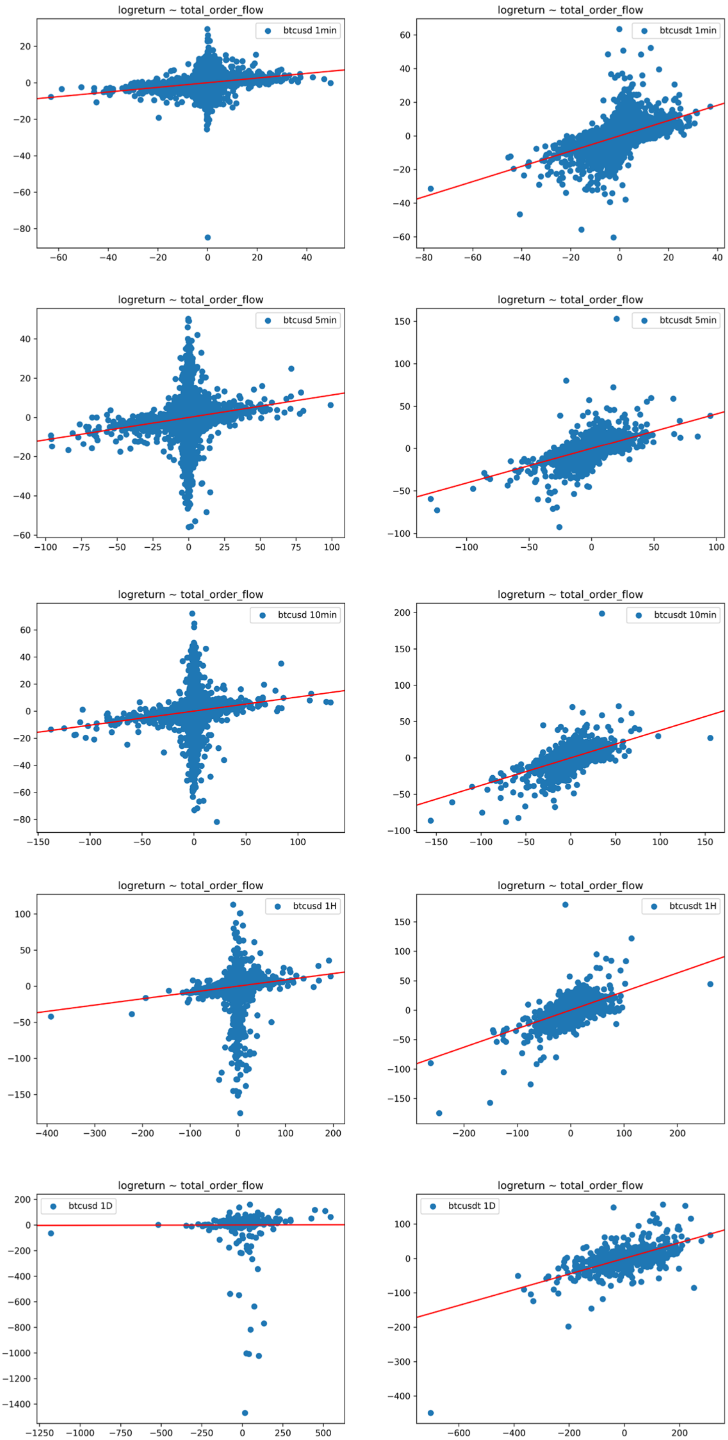

logreturn ~ total_order_flow
btcusd 1min
logreturn ~ total_order_flow
btcusdt 1min
logreturn ~ total_order_flow
btcusd 5min
logreturn ~ total_order_flow
btcusdt 5min
logreturn ~ total_order_flow
btcusd 10min
logreturn ~ total_order_flow
btcusdt 10min
logreturn ~ total_order_flow
btcusd 1H
logreturn ~ total_order_flow
btcusdt 1H
logreturn ~ total_order_flow
btcusd 1D
logreturn ~ total_order_flow
btcusdt 1D

Table 4 OLS regression of r_t on OF_t; OF_t, OF_{t-1}; OF_t, OF_t^3 at 1, 5 and 10 min, hour, daily frequency

Explanatory variable	Log-return														
	1 min			5 min			10 min			1 h			1 day		
	BTC-USD														
Order-Flow$_t$	0.1274	0.1338	0.1345	0.1142	0.1191	0.1171	0.1042	0.1079	0.1060	0.0872	0.0900	0.0836	0.0027	0.0034	− 0.0303
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.9580)	(0.9477)	(0.6369)
Order-Flow$_{t-1}$		− 0.0277			− 0.0251			− 0.0202			− 0.0198			− 0.0874	
		(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(0.0914)	
Order-Flow$_t^3$			-2.5×10^{-5}			-2.2×10^{-6}			-6.2×10^{-7}			2.2×10^{-7}			7.9×10^{-8}
			(< 0.001)			(0.0127)			(0.2647)			(0.1617)			(0.3855)
R^2	1.6226	1.6952	1.6414	2.1761	2.2774	2.1798	2.0559	2.1310	2.0574	1.5684	1.6475	1.5825	0.0005	0.5022	0.1333
	ETH-USD														
Order-Flow$_t$	0.0875	0.0905	0.0973	0.0776	0.0807	0.0883	0.0728	0.0752	0.0788	0.0646	0.0661	0.0665	0.0372	0.0422	0.0368
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.4475)	(0.3889)	(0.5616)
Order-Flow$_{t-1}$		− 0.0144			− 0.0164			− 0.0125			− 0.0104			− 0.0677	
		(< 0.001)			(< 0.001)			(< 0.001)			(0.0846)			(0.1675)	
Order-Flow$_t^3$			-8.8×10^{-6}			-1×10^{-6}			-5.9×10^{-7}			-5.5×10^{-8}			1.1×10^{-9}
			(< 0.001)			(< 0.001)			(< 0.001)			(0.5194)			(0.9939)
R^2	0.7660	0.7858	0.8264	0.9709	1.0126	1.0694	0.9580	0.9853	0.9912	0.8543	0.8759	0.8573	0.1016	0.4370	0.1016
	BTC-USDT														
Order-Flow$_t$	0.4533	0.4655	0.4535	0.4084	0.4251	0.3997	0.3785	0.3963	0.3708	0.3151	0.3295	0.2977	0.2280	0.2634	0.1571
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)

Table 4 (continued)

Explanatory variable	Log-return														
	1 min			5 min			10 min			1 h			1 day		
Order-Flow$_{t-1}$		−0.0710			−0.0851			−0.0819			−0.0703			−0.1072	
		(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)	
Order-Flow$_t^3$			-1.4×10^{-6}			7.3×10^{-6}			3.5×10^{-6}			1.9×10^{-6}			9.6×10^{-7}
			(0.4078)			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)
R^2	20.5462	21.0347	20.5462	24.4600	25.4822	24.5288	25.9916	27.1525	26.0486	30.3742	31.8247	30.7629	35.4542	42.4220	45.7506
	ETH-USDT														
Order-Flow$_t$	0.4966	0.5113	0.5338	0.4222	0.4427	0.4362	0.3857	0.4101	0.3914	0.3041	0.3279	0.2903	0.1676	0.2106	0.1118
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Order-Flow$_{t-1}$		−0.0817			−0.0970			−0.0953			−0.0932			−0.0982	
		(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)	
Order-Flow$_t^3$			-1.7×10^{-4}			-1×10^{-5}			-3.1×10^{-6}			1.3×10^{-6}			4.6×10^{-7}
			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)
R^2	24.6698	25.3163	25.3621	27.0767	28.4410	27.2337	27.6179	29.1899	27.6409	33.0905	35.9903	33.3400	31.2311	39.7967	40.4224
	ETH-BTC														
Order-Flow$_t$	0.2813	0.2833	0.2898	0.2604	0.2646	0.2717	0.2450	0.2515	0.2572	0.1544	0.1790	0.1645	0.0147	0.0931	0.0281
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.1234)	(< 0.001)	(0.1091)
Order-Flow$_{t-1}$		−0.0217			−0.0324			−0.0421			−0.0741			−0.0910	
		(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)	

Table 4 (continued)

Explanatory variable	Log-return														
	1 min			5 min			10 min			1 h			1 day		
Order-Flow$_t^3$			-2.5×10^{-5}			-6×10^{-6}			-6.7×10^{-6}			-4.4×10^{-6}			-1.2×10^{-7}
			(< 0.001)			(< 0.001)			(< 0.001)			(< 0.001)			(0.3629)
R^2	7.9120	7.9584	8.0958	8.6770	8.8092	8.9924	8.7076	8.9583	9.0459	6.0481	7.2890	6.2589	0.4174	4.5200	0.5628
	USDT-USD														
Order-Flow$_t$	−0.0011	−0.0021	−0.0012	0.0005	−0.0013	0.0007	0.0007	−0.0007	0.0010	0.0038	0.0035	0.0031	−0.0052	−0.0027	−0.0112
	(0.4618)	(0.1842)	(0.4957)	(0.8307)	(0.5732)	(0.7807)	(0.7685)	(0.7793)	(0.7413)	(0.3751)	(0.4196)	(0.5348)	(0.6926)	(0.8493)	(0.5273)
Order-Flow$_{t-1}$		0.0030			0.0049			0.0055			0.0010			−0.0077	
		(0.0501)			(0.0282)			(0.0285)			(0.8121)			(0.5821)	
Order-Flow$_t^3$			7.4×10^{-8}			-1.4×10^{-8}			-3.0×10^{-9}			8.3×10^{-9}			4.8×10^{-8}
			(0.9040)			(0.8532)			(0.8773)			(0.8069)			(0.6128)
R^2	0.0002	0.0013	0.0002	0.0000	0.0042	0.0001	0.0001	0.0081	0.0002	0.0060	0.0064	0.0064	0.0275	0.0810	0.0727

Each cell reports the regressor coefficient and the corresponding p-value (below the coefficient, in brackets). R^2 is shown in percentage points

$$r_t = \beta_0 + \beta_1 OF_t + \beta_2 OF_{t-1} + \epsilon_t, \tag{4}$$

with the exception of the regression for ETH-BTC at the daily frequency, where the coefficient of OF_t is positive and statistically significant.

The rationale for the non statistically significant effect detected for the USDT-USD market can be traced back to the features of USDT: being Tether a stablecoin pegged to the US dollar, traders do not attach any informative value to the order flow interpreting it as pure liquidity. Considering BTC-USD and ETH-USD markets, we observe a positive statistically significant coefficient up to 1 h, at the one day frequency the coefficient is not statistically significant. However, the explanatory power of the regression is negligible for all the frequencies.[8] In particular, at the daily frequency we have a very low R^2 and a non significative relationship. Results are different for BTC-USDT and ETH-USDT markets showing high explanatory power of the regressions.[9]

The literature on market impact of OF has investigated the linearity of the relation, see Cont et al. (2013), Silantyev (2019): a large OF should significantly move the price. Fig. 3, the pictures render a visual representation of the results of the regressions. The slope in case of BTC-USD is lower than in case of BTC-USDT, at the daily frequency the relationship for the first market is nearly flat. Moreover, in case of BTC-USD we have many observations with limited OF and a large (in absolute value) log-return, the phenomenon is not observed in the BTC-USDT market. We conclude that prices also move without a significant OF in markets where cryptocurrencies are exchanged with the US dollar, instead price movements are associated with a large OF in markets where cryptocurrencies are exchanged with Tether. As a robustness check we performed a regression eliminating observations with large return/small OF, results look similar to those presented above: the explanatory power of the OF for return in the BTC-USD and ETH-USD market increases in a negligible way.

Confirming the analysis in Silantyev (2019), the pictures show no clear evidence of nonlinearity in the relationship between OF and return. To investigate the point analytically, we have added a cubic term for OF in the regressions (Table 4, third column)

$$r_t = \beta_0 + \beta_1 OF_t + \beta_3 OF_t^3 + \epsilon_t.$$

[8] Results are different from those obtained in Silantyev (2019) for the mid (bid-ask) price changes with a high explanatory power decreasing in the frequency.

[9] The results on the market impact for the BTC-USD market are aligned with those obtained in the literature. Makarov and Schoar (2020) show that the common component of order flow of different exchanges explains a very large fraction of the common component in returns of different exchanges. Regressing the common component in daily returns on the contemporaneous common component in order flow, they compute a price impact of 9 basis points for a 1 million trade. The order flow has a strong explanatory power (R^2 is at 54%). Lyons and Viswanath-Natraj (2019) find lower impact for the US dollar price of Tether (2.5 basis points for a 1 million trade). The price impact is considerably larger than that observed in foreign exchange markets, see (Evans & Lyons, 2002; Berger et al., 2008) (0.05 basis point per 1 million trade), but is smaller than in stock markets, see (Goyenko et al., 2009).

It turns out that the cubic term enters with a small positive and statistically significant coefficient at low frequencies for BTC-USDT and ETH-USDT (with an increase in the explanatory power, especially at the daily frequency), thus showing that a large OF significantly affects price movements in these markets highlighting that the size of the OF amplifies the price variation.

Moving back to regression (4) (Table 4, second column) we observe a negative coefficient for the lagged OF as in Chordia et al. (2002) for all the markets except USDT-USD. The coefficient is statistically significant for BTC-USD and ETH-USD at high frequency but the explanatory power is almost negligible. Lagged OF is statistically significant for all the frequencies for BTC-USDT, ETH-USDT and BTC-ETH markets. As suggested in the above paper for stock exchange, see also Chordia and Subrahmanyam (2004) for a model, this result is consistent with the inventory stabilization hypothesis: sophisticated traders (insiders) have an inventory target for BTC, ETH, USDT and therefore the lagged imbalance is reversed and hence it exerts a negative effect on the contemporaneous return. It is interesting to notice that the phenomenon is observed only in markets where cryptoassets are exchanged with each other and not in markets where BTC and ETH are exchanged with the US dollar. It seems that sophisticated traders mostly trade in the first set of markets with an inventory target, whereas, in markets where cryptocurrencies are exchanged with the US dollar there are outsiders who trade for other reasons.

These results confirm the heterogeneity among crypto-markets: markets where an asset is exchanged with the US dollar and markets where cryptoassets are exchanged between them. In the latter set of markets we observe a strong impact of order flow on market return with a nonlinear effect, moreover in these markets the dynamics of the order flow is consistent with the hypothesis that sophisticated traders operate with an inventory target. In markets where cryptoassets are exchanged with the US dollar the order flow does not seem to contain relevant information.

We may interpret these results as showing that in markets where only cryptoassets are involved (and in particular a stablecoin) traders interpret the order flow as conveying market sentiment-opinions of the market or technology shocks and, therefore, in these markets the price moves in the direction of the order flow. This evidence corroborates the claim that BTC-USDT and ETH-USDT markets play a predominant role in aggregating preference/technology shocks and heterogeneous opinions while the markets where cryptoassets are exchanged against the US dollar play a limited role on price discovery being populated by outsiders who mostly follow the flock.

Table 5 OLS regression of r_t on r_{t-1}; r_{t-1}, OF_{t-1} at 1, 5 min, hour, daily frequency

Explanatory variable	Logreturn														
	1 min			5 min			10 min			1 h			1 day		
	BTC-USD														
Logreturn$_{t-1}$	−0.0313		−0.0323	0.0489		0.0504	0.0845		0.0864	0.3002		0.3063	0.7281		0.7283
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
OrderFlow$_{t-1}$		0.0035	0.0076		−0.0018	−0.0075		−0.0006	−0.0096		−0.0073	−0.0340		−0.0874	−0.0894
		0.0015	(< 0.001)		(0.3521)	(< 0.001)		(0.8246)	(< 0.001)		(0.2232)	(< 0.001)		(0.0912)	(0.0116)
R^2	0.0979	0.0012	0.1036	0.2396	0.0005	0.2488	0.7137	0.0001	0.7307	9.0127	0.0108	9.2468	53.0442	0.5014	53.5692
	ETH-USD														
Logreturn$_{t-1}$	−0.0232		−0.0238	0.0431		0.0436	0.0521		0.0523	0.2039		0.2058	0.6550		0.6573
	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)
OrderFlow$_{t-1}$		0.0047	0.0068		−0.0010	−0.0044		0.0021	−0.0017		−0.0008	−0.0141		−0.0645	−0.0880
		(< 0.001)	(< 0.001)		(0.6087)	(0.0248)		(0.4236)	(0.5067)		(0.8883)	(0.0161)		(0.1869)	(0.0168)
R^2	0.0537	0.0022	0.0582	0.1856	0.0002	0.1887	0.2711	0.0008	0.2717	4.1571	0.0001	4.1976	43.2367	0.3064	43.8063
	BTC-USDT														
Logreturn$_{t-1}$	0.0080		0.0047	−0.0210		−0.0263	−0.0052		−0.0108	−0.0340		−0.0447	−0.1407		−0.1695
	(< 0.001)		0.0001	(< 0.001)		(< 0.001)	(0.1338)		(0.0076)	(< 0.001)		(< 0.001)	(< 0.001)		(0.0011)
OrderFlow$_{t-1}$		0.0093	0.0072		−0.0019	0.0089		0.0041	0.0082		−0.0030	0.0111		−0.0199	0.0185
		(< 0.001)	(< 0.001)		(0.3605)	(< 0.001)		(0.1172)	(0.0067)		(0.5416)	(0.0578)		(0.2153)	(0.3487)
R^2	0.0063	0.0087	0.0104	0.0441	0.0005	0.0529	0.0027	0.0030	0.0117	0.1155	0.0027	0.1418	1.9728	0.2702	2.1246
	ETH-USDT														

Table 5 (continued)

Explanatory variable	Logreturn														
	1 min			5 min			10 min			1 h			1 day		
Logreturn_{t-1}	− 0.0001		− 0.0067	− 0.0143		− 0.0164	− 0.0092		− 0.0225	− 0.0250		− 0.0224	− 0.1301		− 0.1744
	0.9158		(< 0.001)	(< 0.001)		(< 0.001)	(0.0092)		(< 0.001)	(0.0036)		(0.0326)	(0.0019)		(< 0.001)
OrderFlow_{t-1}		0.0099	0.0132		− 0.0036	0.0033		0.0100	0.0186		− 0.0092	− 0.0023		− 0.0052	0.0240
		(< 0.001)	(< 0.001)		(0.0753)	(0.1588)		(< 0.001)	(< 0.001)		(0.0433)	(0.6717)		(0.6831)	(0.1113)
R^2	(< 0.001)	0.0098	0.0132	0.0203	0.0019	0.0215	0.0084	0.0184	0.0549	0.0623	0.0300	0.0636	1.6879	0.0294	2.1271
	ETH-BTC														
Logreturn_{t-1}	0.0025		0.0012	− 0.0035		− 0.0044	− 0.0250		− 0.0262	− 0.0399		− 0.0362	0.0859		0.0892
	(0.0218)		(0.3111)	(0.1605)		(0.0879)	(< 0.001)		(< 0.001)	(< 0.001)		(< 0.001)	(0.0397)		(0.0331)
OrderFlow_{t-1}		0.0052	0.0049		0.0017	0.0028		− 0.0032	0.0032		− 0.0148	− 0.0093		− 0.0108	− 0.0120
		(< 0.001)	(< 0.001)		(0.4403)	(0.2149)		(0.2678)	(0.2908)		(0.0057)	(0.0946)		(0.2602)	(0.2079)
R^2	0.0006	0.0027	0.0028	0.0012	0.0004	0.0021	0.0627	0.0015	0.0641	0.1590	0.0559	0.1794	0.7425	0.2231	1.0200
	USDT-USD														
Logreturn_{t-1}	−0.0261		−0.0261	−0.0071		− 0.0071	0.0214		0.0214	− 0.0010		− 0.0010	− 0.0873		− 0.0877
	(< 0.001)		(< 0.001)	(0.0150)		(0.0151)	(< 0.001)		(< 0.001)	(0.9096)		(0.9069)	(0.0358)		(0.0350)
OrderFlow_{t-1}		0.0024	0.0023		0.0045	0.0045		0.0053	0.0053		0.0019	0.0019		− 0.0086	− 0.0090
		(0.1057)	(0.1104)		(0.0330)	(0.0331)		(0.0283)	(0.0291)		(0.6616)	(0.6609)		(0.5152)	(0.4941)
R^2	0.0679	0.0008	0.0686	0.0051	0.0039	0.0090	0.0461	0.0079	0.0540	0.0001	0.0015	0.0016	0.7735	0.0746	0.8553

Two numbers are reported for each cell: the regressor coefficient and the corresponding p-value which is shown below the coefficient. R^2 is shown in percentage points

5 Market efficiency

In this section we deal with market efficiency investigating the relation between log-return and lagged log-return and OF, that is

$$r_t = \beta_0 + \beta_1 r_{t-1} + \epsilon_t, \tag{5}$$

$$\begin{aligned} r_t &= \beta_0 + \beta_2 OF_{t-1} + \epsilon_t, \\ r_t &= \beta_0 + \beta_1 r_{t-1} + \beta_2 OF_{t-1} + \epsilon_t. \end{aligned} \tag{6}$$

In Table 5 we provide results of the regressions. The evidence is not in favor of market efficiency: the coefficient associated with lagged log-return is statistically significant in almost all the markets at every frequency.

This result is at odds with the evidence for stock markets, where it is shown that the lagged log-return is not statistically significant at any frequency, that lagged OF has an impact on market return intraday over a short horizon, then sophisticated traders react to order imbalances within trading day by undertaking countervailing trades exploiting serial correlation of the OF, see Chordia et al. (2002, 2005).

We concentrate on all markets with the exception of USDT-USD, as it is not really interesting to investigate efficiency in this market being pegged to 1 US dollar. The markets look different along three dimensions: sign of the coefficient of past return, statistical significance of the regression, and statistical significance of the coefficient of OF_{t-1}. BTC-USD and ETH-USD are characterized by a negative coefficient for past return at the one minute frequency and by a positive coefficient at lower frequency. BTC-USDT, ETH-USDT and ETH-BTC markets are characterized by a positive (or non significant) coefficient at one minute frequency and mostly by a negative coefficient for past return at lower frequency. To provide a visual representation of the difference, in Fig. 4 we plot the log-return against lagged log-return for BTC-USD and BTC-USDT, along with the line corresponding to regression (5). For the other pairs, we refer to Figs. 8, 9 in Appendix 1. This evidence suggests that the first couple of markets is inefficient with the predominance of trend-follower traders during the day, instead the second set of markets is characterized by contrarian traders. This result corroborates the evidence on the OF in Sect. 3. Looking at the explanatory power of the regressions, we observe that it is negligible at a high frequency but becomes very high at the daily frequency in BTC-USD and ETH-USD markets. Instead, in BTC-USDT, ETH-USDT and ETH-BTC markets the explanatory power is negligible and increases slightly at the daily frequency. As far as the lagged OF is concerned, looking at regression (6), we observe that its coefficient is statistically significant for BTC-USD and ETH-USD at any frequency (with only one exception). Instead, in BTC-USDT, ETH-USDT and ETH-BTC markets the coefficient is statistically significant only at high frequency and not at low frequency. In both cases, the contribution to the explanatory power is limited.

We should be cautious in interpreting these results as the explanatory power is very low with the only exception provided by BTC-USD and ETH-USD at daily frequency. The two sets of markets seem to be inefficient but in a different way. In case of markets where cryptocurrencies are exchanged with the US dollar, predictability

Fig. 4 Scatter plot of BTC-USD (left) and BTC-USDT (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: lagged log-return, *y*-axis: log-return. The straight line obtained from the linear regression is reported ▶

comes from a trend component in returns which is likely to be associated with a herding phenomenon among traders, see Ballis and Drakos (2020), Bouri et al. (2019), King & Koutmos (2021), Manahov (2021); also lagged OF has a predictive power for future return. At daily horizon, predictability is high and the evidence is strongly against markets efficiency. Instead, in markets where cryptoassets are exchanged with each other, there is evidence of mean reversion and predictability is limited. This result may be linked to liquidity effects rather than to traders exploiting predictability. However, the results at a daily frequency suggest that markets where cryptocurrencies are exchanged with the US dollar are strongly inefficient.

6 Arbitrage profits

The emergence of arbitrage opportunities in the BTC-USD market has been investigated in Makarov and Schoar (2020) considering different exchanges and different currencies, i.e., an arbitrage is obtained buying/selling BTC in different exchanges for different currencies. They find out that there are limited arbitrage opportunities in each market (exchanges per a specific currency) but significant arbitrage opportunities arise by trading in different currencies. In what follows, we apply their methodology as it is described in Appendix 1. We define an arbitrage strategy as a couple of trades that can be implemented in a market buying and selling the cryptoassets contemporaneously (same second as a time stamp) with no inventory risk. For example an arbitrage strategy in the BTC-USD market is built acquiring BTC (with USD) in an exchange and immediately reselling them in another exchange obtaining a positive net amount of USD. To implement this strategy, the price of the first transaction should be lower than the one of the second transaction yielding an arbitrage spread ($s_{\text{ARBITRAGE}}$, see Eq. (7) in Appendix 1). The profit of the arbitrage strategy ($p_{\text{ARBITRAGE}}$, see Eq. (8)) is obtained by multiplying the spread for the minimum between the quantity available for trade on the bid and on the ask side.

Table 6 provides summary statistics for the arbitrage strategy. Confirming the analysis in Makarov and Schoar (2020), the money value of arbitrage opportunities (*arbitrage*) is rather limited. The rank of arbitrage profits presents on top BTC-USDT, ETH-USDT and BTC-USD markets, the market with the smallest amount of arbitrage profit is the one for USDT-USD. We decompose arbitrage profits in the average arbitrage size (*spread*) and the fraction of seconds in the sample with an arbitrage opportunity (*opp_perc*). The first measure is about the size of the arbitrage when it materializes, the second one about its frequency. The average arbitrage size is high in markets where BTC, ETH and USDT are exchanged with the USD, but in these markets there are few seconds with arbitrage opportunities. Instead, the frequency of arbitrage opportunities is high in markets where BTC and ETH are exchanged with the USDT but in these markets the average arbitrage spread is limited. Notice that there is a weak connection between the size of the arbitrage opportunity, its frequency and the number of exchanges in which the pair is traded.

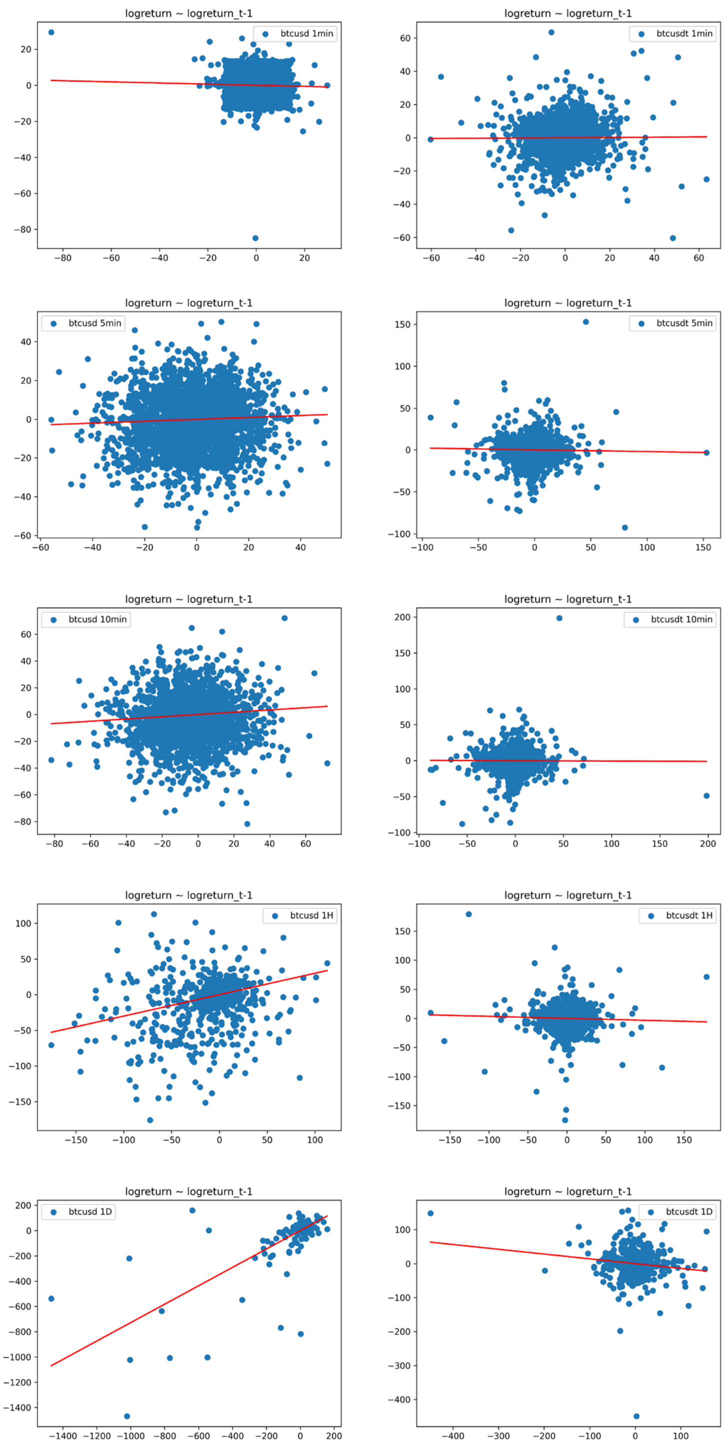

logreturn ~ logreturn_t-1
btcusd 1min
logreturn ~ logreturn_t-1
btcusdt 1min
logreturn ~ logreturn_t-1
btcusd 5min
logreturn ~ logreturn_t-1
btcusdt 5min
logreturn ~ logreturn_t-1
btcusd 10min
logreturn ~ logreturn_t-1
btcusdt 10min
logreturn ~ logreturn_t-1
btcusd 1H
logreturn ~ logreturn_t-1
btcusdt 1H
logreturn ~ logreturn_t-1
btcusd 1D
logreturn ~ logreturn_t-1
btcusdt 1D

Table 6 Arbitrage summary statistics trading in a single pair

Market	Volume	Arbitrage	Opp	Opp_perc (%)	Spread	#exchanges	Net profits
BTC-USDT	1.87 × 109	6.20 × 106	9,725,039	20.84%	4.73	7	0.00
ETH-USDT	4.79 × 108	2.28 × 106	6,129,799	13.14%	4.64	6	0.00
ETH-BTC	2.58 × 108	3.22 × 105	2,206,479	4.73%	5.70	15	0.00
BTC-USD	3.71 × 108	2.86 × 106	953,282	2.04%	26.05	10	665464.04
ETH-USD	9.57 × 107	5.63 × 105	220,560	0.47%	41.75	10	293533.80
USDT-USD	1.24 × 107	6.82 × 103	6034	0.01%	19.96	2	0.00

Trading volume (*volume*) is the mean daily trading volume expressed in USD; *arbitrage* represents the total profits in USD generated by trading exploiting arbitrage opportunities (sum of $p_{\text{ARBITRAGE}}$); arbitrage opportunity (*opp*) is the number of 1 s intervals with a strictly positive arbitrage opportunity; *opp_perc* is the percentage of seconds with a strictly positive arbitrage opportunity; *spread* provides the size of the arbitrage opportunity averaging for each 1 s interval with a strictly positive arbitrage opportunity ($s_{\text{ARBITRAGE}} > 0$), the spread is computed in basis points with respect to the corresponding bid-ask mid price. We also report the number of exchanges for each pair (see Appendix 1), and the arbitrage profit net of transaction costs (*net_profits*)

Most exchanges have a taker fee of 0.10%.[10] In the last column of Table 6 we report arbitrage profits net of transaction costs. High fees may render unprofitable arbitrage opportunities. This occurs in case of the pair with highest arbitrage profits (BTC-USDT). We can conclude that arbitrage strategies between USDT and cryptocurrencies are non-profitable. Instead, arbitrage strategies centered on USD (BTC-USD, ETH-USD) are profitable net of transaction costs. The outcome is due to the fact that latter markets are characterized by higher arbitrage spreads.

We have extended the analysis to arbitrage opportunities that can arise when three markets are involved, considering a cryptocurrency, Tether and the US dollar, see Fig. 5 in Appendix 1 for a graphical illustration. Table 7 provides summary statistics for the *buy* and *sell* arbitrage strategies: considering a triangulation, a buy arbitrage strategy consists in buying Tether through the BTC or ETH markets, and selling them in the Tether market against USD; a sell arbitrage strategy goes in the opposite direction, see Appendix 1 for details. We observe that the size of arbitrage profits is rather limited. Arbitrage profits net of fees are null in most of the cases. Because of the limited size of the arbitrage we omit further analysis.

[10] Crypto exchanges adopt a maker-taker fee schedule based on the rolling 30-day cumulative trading volume. The majority of exchanges apply the same schedule to all pairs. Some of them differentiate among pairs. The maker fee is equal or less than the taker fee in order to incentivize traders to provide liquidity. Both fees decrease in trading volume. The maker fee is paid by the trader posting the quote to the order book while the taker fee is paid by the trader filling the quote and initiating the trade. For the pairs considered in Table 6, we have collected the current taker fees for the first level of the schedules of all exchanges trading the corresponding pair. In this way we are using the highest possible fee for executing the trades and thus we are conservative in profit estimates (likely actual fees were lower). Unfortunately, we do not have access to the historical fee schedules during our period of analysis. Nevertheless, the current fee levels are representative of transaction costs during the sample. Table 10 in Appendix 1 shows the fees for the different pairs.

Table 7 Arbitrage summary statistics for the *buy* and *sell* arbitrage strategies, trading in three pairs

Markets	Arbitrage	Opp	Spread
Buy			
BTC-USDT, BTC-USD, USDT-USD	262,706.85	219,367	27.69
ETH-USDT, ETH-USD, USDT-USD	162,793.55	135,006	30.87
Sell			
BTC-USDT, BTC-USD, USDT-USD	362,011.00	349,797	21.43
ETH-USDT, ETH-USD, USDT-USD	167,245.30	224,373	29.40

Arbitrage represents the profits in US dollars generated by trading exploiting arbitrage opportunities (summation of $p^j_{\text{ARBITRAGE}}$, $j = \text{SELL,BUY}$, see Eq. (11) in Appendix 1); arbitrage opportunities (*opp*) are the number of 1 s intervals with a strictly positive arbitrage opportunity; *spread* provides the size of the arbitrage opportunity for each 1 s interval with a strictly positive arbitrage opportunity ($s^j_{\text{ARBITRAGE}} > 0$, $j = \text{SELL,BUY}$, see Eqs. (9)–(10) in Appendix 1), the spread is computed in basis points with the respect to the corresponding bid-ask mid price

We investigate efficiency in crypto-markets by looking at the relationship between the size of the arbitrage spread and market activity: trading volume and the absolute value of OF. If markets are efficient, we expect market activity to close arbitrage opportunities. Notice that the arbitrage spread is positive or null and therefore we provide censored regressions. For the USDT-USD market, given that the arbitrage spread is different from zero only on 0.01% of the seconds, see Table 6, we substitute the arbitrage spread with the price parity, i.e., the distance of the value in USD of Tether from 1 USD. We control for the lagged level of the arbitrage spread, and therefore our regression looks at the arbitrage spread variation.

There are few papers dealing with the relation between arbitrage opportunities and market activity in stock exchange markets,[11] in our analysis we deal with the following censored regressions

$$s_{\text{ARBITRAGE }t} = \beta_0 + \beta_1 s_{\text{ARBITRAGE }t-1} + \beta_2 |OF_t| + \epsilon_t,$$
$$s_{\text{ARBITRAGE }t} = \beta_0 + \beta_1 s_{\text{ARBITRAGE }t-1} + \beta_2 V_t + \epsilon_t,$$

V denoting the volume. In Chordia et al. (2002), the authors analyze the relationship between the absolute value of OF and variations in bid-ask spread (liquidity). They find that higher bid-ask spreads, and therefore lower arbitrage spreads, occur when orders are more unbalanced in both directions. Building on this result, we claim that $|OF|$ should negatively affect the arbitrage size: order imbalance in either direction should close the arbitrage opportunities. Results reported in the left part of Table 8 show that this is not the case in crypto-markets: a significant order imbalance (from the buy and sell side) leads to a larger arbitrage spread. A similar result holds true for trading volume, see the right part of Table 8. Only in case of USDT-USD at a daily frequency the variables are not statistically significant (also USDT-USD at one minute frequency for $|OF|$).

[11] The paper closest to our in the spirit is Roll et al. (2007). They show that liquidity (computed as bid-ask spread) enhances efficiency in the future-cash pricing system reducing the futures-cash basis.

Table 8 Censored regressions of the arbitrage spread on market activity at 1, 5 and 10 min, hour, daily frequency

Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)					Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)				
	1 min	5 min	10 min	1 h	1 day		1 min	5 min	10 min	1 h	1 day
	BTC-USD						*BTC-USD*				
$s_{\text{ARBITRAGE } t-1}$	0.8280	0.8511	0.8246	0.8384	0.7894	$s_{\text{ARBITRAGE } t-1}$	0.7501	0.7226	0.6502	0.6039	0.4527
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.3669	0.1553	0.1172	0.0380	0.0036	Volume$_t$	0.4316	0.1252	0.07828	0.0161	0.0008
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
R^2	53.8842	57.9612	52.5772	51.0229	48.0784	R^2	54.5485	58.8021	54.1702	56.9928	57.1924
	ETH-USD						*ETH-USD*				
$s_{\text{ARBITRAGE } t-1}$	0.6391	0.7166	0.6657	0.6748	0.7763	$s_{\text{ARBITRAGE } t-1}$	0.4946	0.5011	0.4005	0.3439	0.3704
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.5993	0.1495	0.1022	0.0262	0.0024	Volume$_t$	0.7551	0.1368	0.0785	0.0139	0.0005
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
R^2	25.4324	41.5112	38.0920	37.8995	46.0226	R^2	28.9296	47.8862	47.16321	55.8521	68.0196
	BTC-USDT						*BTC-USDT*				
$s_{\text{ARBITRAGE } t-1}$	0.9709	0.9527	0.9300	0.8102	0.9256	$s_{\text{ARBITRAGE } t-1}$	0.9147	0.8721	0.8323	0.6542	0.7024
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.2192	0.1080	0.0818	0.0359	0.0030	Volume$_t$	0.2619	0.0783	0.0484	0.0135	0.0003
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
R^2	77.3691	73.2160	67.4819	43.0079	46.3419	R^2	76.8923	72.0387	66.0954	45.1223	47.8793
	ETH-USDT						*ETH-USDT*				
$s_{\text{ARBITRAGE } t-1}$	0.9792	0.9538	0.9091	0.7393	0.8460	$s_{\text{ARBITRAGE } t-1}$	0.9260	0.8651	0.7866	0.5102	0.4083
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.2121	0.1103	0.0898	0.0335	0.0020	Volume$_t$	0.2581	0.0892	0.0623	0.0164	0.0005
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
R^2	85.1020	77.9328	66.3602	34.5846	30.3748	R^2	84.3140	76.3001	64.7002	40.4105	46.8280

Table 8 (continued)

Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)					Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)				
	1 min	5 min	10 min	1 h	1 day		1 min	5 min	10 min	1 h	1 day
	ETH-BTC						*ETH-BTC*				
$s_{\text{ARBITRAGE } t-1}$	0.7855	0.9281	0.9230	0.8945	0.8855	$s_{\text{ARBITRAGE } t-1}$	0.6684	0.7684	0.7277	0.6160	0.3431
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.3071	0.1019	0.0760	0.0296	0.0016	Volume$_t$	0.4603	0.0949	0.0566	0.0134	0.0006
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
R^2	40.2842	63.0422	61.5206	54.5535	41.3046	R^2	42.0026	64.6975	64.2116	63.7371	64.2036
	USDT-USD						*USDT-USD*				
PriceParity$_{t-1}$	0.8410	0.8035	0.8477	0.8503	0.9046	PriceParity$_{t-1}$	0.8410	0.8035	0.8476	0.8503	0.9044
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\|\text{OrderFlow}\|_t$	0.0015	0.0021	0.0020	0.0009	− 0.0000	Volume$_t$	0.0020	0.0020	0.0017	0.0007	0.0000
	(0.0718)	(0.0022)	(< 0.001)	(0.0212)	(0.7852)		(0.0128)	(0.0022)	(< 0.001)	(0.0301)	(0.7526)
R^2	70.8405	64.3969	72.6671	72.7039	82.5651	R^2	70.8408	64.3968	72.6667	72.7027	82.5659

Each cell reports the regressor coefficient and the corresponding p-value (below the coefficient, in brackets). R^2 is shown in percentage points

Table 9 Censored regressions of the arbitrage spread on OF_t and OF_t OF_{t-1} at 1, 5 and 10 min, hour, daily frequency

Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)									
	1 min	5 min	10 min	1 h	1 day	1 min	5 min	10 min	1 h	1 day
	BTC-USD									
$s_{\text{ARBITRAGE } t-1}$	0.8973	0.9535	0.9396	0.9720	1.0145	0.8991	0.9511	0.9374	0.9715	1.0098
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	0.0662	0.0028	− 0.0037	− 0.0021	− 0.0001	0.0882	0.0080	− 0.0021	− 0.0020	− 0.0001
	(< 0.001)	(0.08105)	(0.0170)	(0.04890)	(0.7035)	(< 0.001)	(< 0.001)	(0.1772)	(0.0575)	(0.7057)
OrderFlow$_{t-1}$						− 0.08978	− 0.0253	− 0.0083	− 0.0003	− 0.0003
						(< 0.001)	(< 0.001)	(< 0.001)	(0.6949)	(0.3153)
R^2	53.7906	56.4627	48.7937	45.0736	28.3740	54.2234	56.9361	48.9871	45.1003	29.1305
	ETH-USD									
$s_{\text{ARBITRAGE } t-1}$	0.7556	0.8244	0.7818	0.7812	0.8721	0.7537	0.8178	0.7761	0.7784	0.9171
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	0.0662	− 0.0141	− 0.0158	− 0.0053	− 0.0008	0.0952	− 0.0105	− 0.0144	− 0.0052	− 0.0008
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.0011)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_{t-1}$						− 0.1338	− 0.0183	− 0.0070	− 0.0006	0.0004
						(< 0.001)	(< 0.001)	(< 0.001)	(0.3965)	(0.1093)
R^2	23.8196	37.9154	32.3844	31.0113	33.5122	24.3959	38.2671	32.5970	31.0869	32.2170
	BTC-USDT									
$s_{\text{ARBITRAGE } t-1}$	1.0009	0.9938	0.9784	0.8723	0.9662	0.9999	0.9932	0.9779	0.8669	0.9952
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	0.0273	− 0.0069	− 0.0083	− 0.0051	− 0.0002	0.0349	− 0.0062	− 0.0079	− 0.0045	− 0.0005
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.4548)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.0717)
OrderFlow$_{t-1}$						− 0.0396	− 0.0035	− 0.0017	− 0.0029	0.0011
						(< 0.001)	(0.0013)	(0.1015)	(0.0050)	(< 0.001)
R^2	77.1316	72.1640	65.7788	38.1092	40.1180	77.2316	72.1825	65.7972	38.4381	40.6095

Table 9 (continued)

Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)									
	1 min	5 min	10 min	1 h	1 day	1 min	5 min	10 min	1 h	1 day
	ETH-USDT									
$s_{\text{ARBITRAGE } t-1}$	1.0092	0.9934	0.9545	0.7968	0.8702	1.0085	0.9929	0.9538	0.8007	0.8871
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	0.0046	− 0.0116	− 0.0118	− 0.0055	0.0002	0.0096	− 0.0109	− 0.0112	− 0.0060	− 0.0001
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.2952)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.6277)
OrderFlow$_{t-1}$						− 0.0252	− 0.0033	− 0.0021	0.0019	0.0010
						(< 0.001)	(0.0014)	(0.0735)	(0.0537)	(< 0.001)
R^2	85.2513	77.7737	66.1591	31.4662	22.4821	85.2920	77.7889	66.1866	31.2684	23.2503
	ETH-BTC									
$s_{\text{ARBITRAGE } t-1}$	0.8227	0.9687	0.9688	0.9382	0.8395	0.8245	0.9682	0.9674	0.9319	0.8391
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	0.1154	0.0284	0.0192	0.0119	0.0011	0.1196	0.0274	0.0172	0.0097	0.0009
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.0221)
OrderFlow$_{t-1}$						− 0.04608	0.0076	0.0135	0.0068	0.0002
						(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.5214)
R^2	40.6783	63.5646	61.8016	53.6395	40.8154	40.8599	63.4774	61.5388	52.8390	40.5521
	USDT-USD									
PriceParity$_{t-1}$	0.8407	0.8029	0.8469	0.8481	0.8927	0.8405	0.8027	0.8466	0.8474	0.8919
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_t$	− 0.0079	− 0.0059	− 0.0041	− 0.0031	− 0.0012	− 0.0066	− 0.0050	− 0.0037	− 0.0029	− 0.0012
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
OrderFlow$_{t-1}$						− 0.0037	−0.0027	−0.0016	−0.0008	− 0.0000
						(< 0.001)	(< 0.001)	(0.0013)	(0.0432)	(0.7958)

Table 9 (continued)

Explanatory variable	$s_{\text{ARBITRAGE } t}$ (PriceParity$_t$ for USDT-USD)									
	1 min	5 min	10 min	1 h	1 day	1 min	5 min	10 min	1 h	1 day
R^2	70.8484	64.4160	72.6911	72.8212	84.0106	70.8501	64.4201	72.6958	72.8297	84.0125

Each cell reports the regressor coefficient and the corresponding p-value (below the coefficient, in brackets). R^2 is shown in percentage points

Results change when the the arbitrage spread is regressed on OF, that is

$$s_{\text{ARBITRAGE}\ t} = \beta_0 + \beta_1 s_{\text{ARBITRAGE}\ t-1} + \beta_2 OF_t + \epsilon_t,$$

see the left part of Table 9. If there is a strong activity from the buy side compared to the sell side, then the arbitrage spread declines. The result suggests that traders, aiming at exploiting an arbitrage opportunity, opt to buy the cryptoassets and then to sell rather than to do the reverse. The rationale could be that there are short sale constraints in the markets.[12] The result holds true for all the markets depending on the frequency, except ETH-BTC which shows a statistically significant positive coefficient of OF_t at every frequency. For BTC-USD, ETH-USD, BTC-USDT, ETH-USDT markets we observe a positive statistically significant coefficient for OF_t at one minute, then the coefficient becomes statistically significant and negative and finally tends to be non significant at the daily frequency (the coefficient is still positive at 5 min for BTC-USD, while for ETH-USD is statistically significant also at the daily frequency). This result suggests that markets are not able to close the arbitrage in the short term, i.e., there is a continuation/amplification effect, then traders discover arbitrage opportunities and the market moves in the direction of closing them.[13]

The analysis is similar for the different markets, however the explanatory power of the regressions for BTC-USDT and ETH-USDT is higher than for the regressions for BTC-USD and ETH-USD at high frequency. The first set of markets seems to close arbitrage opportunities quickly compared to the second one. The result is confirmed by observing that coefficient for OF_t is positive also at the five minute frequency for BTC-USD. Once again, these results reinforce the earlier findings, indicating that all markets exhibit inefficiencies. However, the markets where a cryptoasset is exchanged with the US dollar demonstrate particularly strong inefficiency. As a robustness check, in the right part of Table 9 we also consider the lagged OF in the regression, i.e.,

$$s_{\text{ARBITRAGE}\ t} = \beta_0 + \beta_1 s_{\text{ARBITRAGE}\ t-1} + \beta_2 OF_t + \beta_3 OF_{t-1} + \epsilon_t.$$

Results on the coefficient of OF_t are confirmed.

[12] Exchanges apply different policies and different limitations for short selling and, in general, for margin trading. As examples, see https://support.kraken.com/hc/en-us/articles/4402532394260 regarding limitations and https://support.kraken.com/hc/en-us/articles/204585998-Collateral-currency regarding accepted collateral currencies on Kraken; Coinbase Pro started to accept margin trading only in February 2020 https://blog.coinbase.com/margin-trading-is-now-available-on-coinbase-pro-b22743a0e07b, but the feature was disabled in November 2020.

[13] This result is coherent with the analysis in Marshall et al. (2013) on ETF arbitrage: order imbalance (absolute value) increase significantly in minutes surrounding an arbitrage opportunities with markets becoming more one-sided.

7 Conclusions

This paper aims to investigate the microstructure of cryptoasset markets building on the observation that markets where cryptocurrencies are exchanged with the US dollar are different from markets where cryptoassets are exchanged with each other. As a matter of fact, stablecoins allow traders to trade cryptocurrencies with lower costs compared to go through an exchange in US dollars. As a consequence, sophisticated traders are likely to remain inside the cryptoassets domain detaining Tether as safe asset rather than US dollar.

Investigating market impact and efficiency at different frequencies we have provided evidence that markets where cryptoassets are exchanged with each other play a central role on price formation. In these markets there are sophisticated traders who behave as contrarians, instead in markets where cryptoassets are exchanged with the US dollar there is a predominance of herding. In markets where cryptoassets are traded against US dollar the order flow does not seem to contain relevant pieces of information. Moreover, markets where cryptocurrencies are traded against the US dollar are strongly inefficient, whereas those where cryptoassets are exchanged with each other are inefficient to a less extent.

These results highlight that crypto-markets are not homogeneous. To capture the sentiment about cryptoassets, we should not look at markets where cryptocurrencies are exchanged with the US dollar, but at Tether that plays a central role in the cryptoasset environment. Tether–Bitcoin and Tether–Ether are the markets to look at in order to capture the mood about cryptoassets.

Appendix 1: Exchanges

See Table 10.

Table 10 The 6 pairs considered in the analysis with the corresponding exchanges

Pair	Exchanges	Fees (%)
ETH-BTC	Binance	0.10
	HitBTC	0.09
	Huobi Global	0.09
	YoBit	0.20
	OKEx	0.10
	BTC-Alpha	0.10
	KuCoin	0.10
	Coinbase Pro	0.10
	Poloniex	0.13
	Bittrex	0.10
	Bitfinex	0.10
	Bibox	0.10
	BeQuant	0.10
	BigONE	0.06
	Kraken	0.09
BTC-USD	Coinbase Pro	0.10
	Bitstamp	0.10
	Kraken	0.09
	Bitfinex	0.10
	FTX	0.20
	Gemini	0.20
	BTC-Alpha	0.10
	OKCoin	0.10
	HitBTC	0.09
	Huobi Global	0.09
ETH-USD	Coinbase Pro	0.10
	Bitstamp	0.10
	Kraken	0.09
	Bitfinex	0.10
	FTX	0.20
	Gemini	0.20
	BTC-Alpha	0.10
	OKCoin	0.10
	HitBTC	0.09
	YoBit	0.20
BTC-USDT	Binance	0.10
	OKEx	0.10
	Huobi Global	0.09
	BeQuant	0.10
	Bibox	0.10
	BigONE	0.06
	KuCoin	0.10

Table 10 (continued)

Pair	Exchanges	Fees (%)
ETH-USDT	Binance	0.10
	CoinEX	0.20
	OKEx	0.10
	Huobi Global	0.09
	BigONE	0.06
	KuCoin	0.10
USDT-USD	Kraken	0.09
	Bitfinex	0.10

In the last column we report the highest taker fees

Appendix 2: Arbitrage opportunities

In what follows we show how to compute arbitrage opportunities in the markets. The analysis is similar to the one developed in Makarov and Schoar (2020).

First of all, we consider arbitrage opportunities built trading a single pair across different exchanges. We refer to a 1 s time interval. For each 1 s time interval we proceed as follows:

1. For each exchange x, we calculate P_x^{BID} as the size-weighted average of buyer-initiated trades. Analogously, we calculate P_x^{ASK} as the size-weighted average of seller-initiated trades. We keep track of the volume associated to each price.
2. We calculate $P^{\text{ASK}} := \min_x P_x^{\text{ASK}}$ and $P^{\text{BID}} := \max_x P_x^{\text{BID}}$, i.e., the best ask and the best bid price across all the exchanges. We keep track of the exchange and volume associated to P^{ASK}, P^{BID}.
3. For each interval, we calculate the arbitrage spread as
$$s_{\text{ARBITRAGE}} := (P^{\text{BID}} - P^{\text{ASK}})^+. \tag{7}$$
4. If it is not possible to compute one of the two values, then we set the arbitrage spread to zero.
5. We compute the arbitrage profit $p_{\text{ARBITRAGE}}$ as
$$p_{\text{ARBITRAGE}} := s_{\text{ARBITRAGE}} \cdot \min\left[\text{Vol}\,(P^{\text{BID}}),\ \text{Vol}\,(P^{\text{ASK}})\right], \tag{8}$$
where $\text{Vol}\,(P)$ denotes the volume associated to price P. We observe that the arbitrage profit is denominated in the reference currency.

$p_{\text{ARBITRAGE}}$ represents an estimate of the profits that can be obtained in the market exploiting arbitrage opportunities. It is an estimate from below as we match only the best bid and the best ask offer. For example, if we have the following order book $(P^{\text{BID}}, Vol(P^{\text{BID}})) = (2, 10)$, $(P^{\text{ASK}}, Vol(P^{\text{ASK}})) = \{(1, 4), (1.2, 6)\}$, being the best ask price 1, we estimate the arbitrage profits as $(2 - 1) \cdot \min[4, 10] = 4$, however the real arbitrage could be $4 \cdot (2 - 1) + 6 \cdot (2 - 1.2)$ exploiting all the order book.

We extend the approach to arbitrage opportunities that can arise when three markets are involved. Figure 5 shows the triangulation for BTC-USDT-USD. Three markets are involved by the trading strategy: BTC-USDT, BTC-USD, and USDT-USD. To quantify the profits of an arbitrage opportunity we proceed as follows for each 1 s time interval:

1. for BTC-USDT, we calculate $P^{\text{ASK}}_{\text{BTC-USDT}} := \min_x P^{\text{ASK}}_x$, i.e., the best price to buy BTC with USDT.
2. For BTC-USD, we calculate $P^{\text{BID}}_{\text{BTC-USD}} := \max_x P^{\text{BID}}_x$, i.e., the best price to sell BTC for USD.
3. For USDT-USD and for each time interval, we calculate $P^{\text{ASK}}_{\text{USDT-USD}} := \min_x P^{\text{ASK}}_x$, i.e., the best price to buy USDT with USD.
4. the arbitrage spread $s^{\text{SELL}}_{\text{ARBITRAGE}}$ is given by

$$s^{\text{SELL}}_{\text{ARBITRAGE}} := \left(\frac{P^{\text{BID}}_{\text{BTC-USD}}}{P^{\text{ASK}}_{\text{BTC-USDT}}} - P^{\text{ASK}}_{\text{USDT-USD}} \right)^{+}. \tag{9}$$

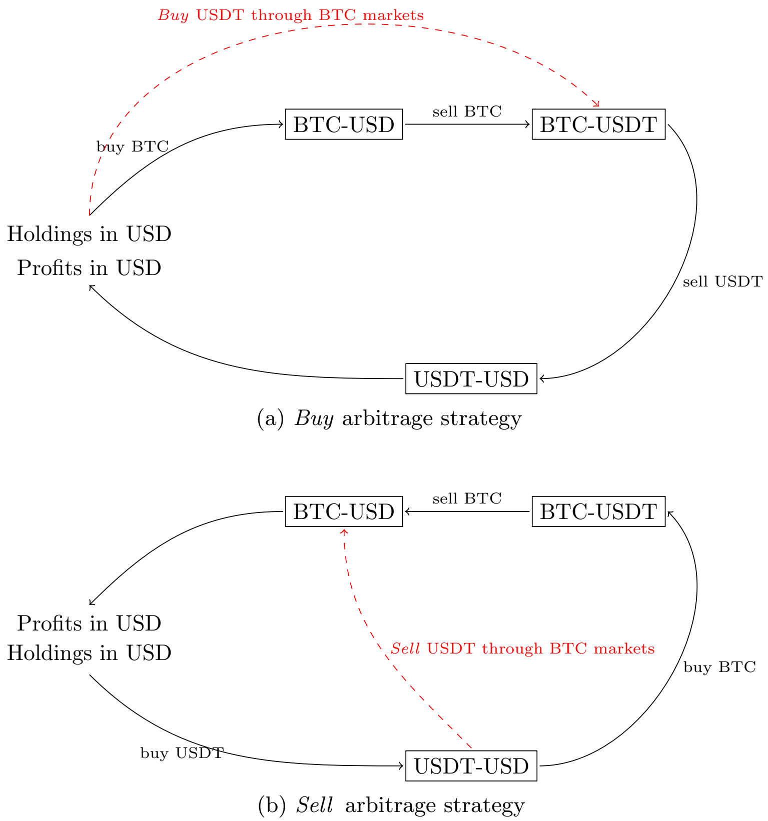


Fig. 5 Triangulations Tether, Bitcoin and US Dollar

this arbitrage is built buying USDT with USD in the USDT-USD market and selling them in the BTC-USDT and BTC-USD markets.

5. Analogously, we define $s^{\text{BUY}}_{\text{ARBITRAGE}}$ as follows,

$$s^{\text{BUY}}_{\text{ARBITRAGE}} := \left(P^{\text{BID}}_{\text{USDT-USD}} - \frac{P^{\text{ASK}}_{\text{BTC-USD}}}{P^{\text{BID}}_{\text{BTC-USDT}}} \right)^{+}. \tag{10}$$

6. The trading volume allowed by the arbitrage opportunity is denoted by $vol_{\text{ARBITRAGE}}$. We denote by $\text{Vol}\,(P^{\text{ASK}}_{\text{BTC-USDT}})$, $\text{Vol}\,(P^{\text{BID}}_{\text{BTC-USD}})$ and $\text{Vol}\,(P^{\text{ASK}}_{\text{USDT-USD}})$ the trading volume associated to the prices defining the above arbitrage spreads. So each spread (for the three markets) is associated to its trading volume. The first two are denominated in BTC and the last one in USDT. Denoting by $P^{\text{AVG}}_{\text{BTC-USDT}}$ the size-weighted average trading price for the pair during the interval, we obtain

$$\begin{aligned} vol^{\text{SELL}}_{\text{ARBITRAGE}} = \min \big[&\text{Vol}\,(P^{\text{ASK}}_{\text{BTC-USDT}}) \cdot P^{\text{AVG}}_{\text{BTC-USDT}}, \\ &\text{Vol}\,(P^{\text{BID}}_{\text{BTC-USD}}) \cdot P^{\text{AVG}}_{\text{BTC-USDT}}, \text{Vol}\,(P^{\text{ASK}}_{\text{USDT-USD}})\big]. \end{aligned}$$

Similarly, we define

$$\begin{aligned} vol^{\text{BUY}}_{\text{ARBITRAGE}} = \min \big[&\text{Vol}\,(P^{\text{BID}}_{\text{BTC-USDT}}) \cdot P^{\text{AVG}}_{\text{BTC-USDT}}, \text{Vol} \\ &(P^{\text{ASK}}_{\text{BTC-USD}}) \cdot P^{\text{AVG}}_{\text{BTC-USDT}}, \text{Vol}\,(P^{\text{BID}}_{\text{USDT-USD}})\big]. \end{aligned}$$

Notice that $vol^{j}_{\text{ARBITRAGE}}$, $j = \text{SELL,BUY}$. is denominated in USDT.

7. The arbitrage profit in USD can be computed as

$$p^{j}_{\text{ARBITRAGE}} := s^{j}_{\text{ARBITRAGE}} \cdot vol^{j}_{\text{ARBITRAGE}}, \quad j = \text{SELL,BUY}. \tag{11}$$

Considering a triangulation, there are two different arbitrage opportunities: a *buy arbitrage* that consists in buying an amount of stablecoin going through the BTC or ETH markets, and selling it in the stablecoin market against USD; a *sell arbitrage* that goes in the opposite direction.

Appendix 3. Autocorrelations of time series

See Table 11.

Table 11 Autocorrelations of market quantities at 5 min, 1 h and 1 day frequencies

Symbol	Interval	Measure	t − 1	t − 2	t − 3	t − 4	t − 5
BTC-USD	5 min	Order Flow	0.1960	0.0888	0.0635	0.0455	0.0385
BTC-USD	5 min	Log-return	0.0489	0.0361	0.0464	0.0309	0.0259
BTC-USD	5 min	Volatility	0.5175	0.4970	0.4975	0.4931	0.4828
BTC-USD	5 min	Arbitrage spread	0.7516	0.5952	0.5153	0.4700	0.4444
BTC-USD	1 H	Order flow	0.1393	0.0792	0.0664	0.0362	0.0371
BTC-USD	1 H	Log-return	0.3002	0.3207	0.3249	0.2606	0.2913
BTC-USD	1 H	Volatility	0.9009	0.8717	0.8546	0.8501	0.8364
BTC-USD	1 H	Arbitrage Spread	0.6700	0.4692	0.3948	0.3094	0.2477
BTC-USD	1 D	Order flow	**0.0076**	0.0984	**0.0813**	0.2003	0.0852
BTC-USD	1 D	Log-return	0.7281	0.4998	0.5126	0.4766	0.4229
BTC-USD	1 D	Volatility	0.9406	0.8729	0.7896	0.7130	0.6182
BTC-USD	1 D	Arbitrage spread	0.5322	0.2341	0.2731	0.3363	0.2715
ETH-USD	5 min	Order flow	0.1907	0.0871	0.0698	0.0384	0.0484
ETH-USD	5 min	Log-return	0.0431	0.0254	0.0431	0.0273	0.0240
ETH-USD	5 min	Volatility	0.6715	0.6521	0.6499	0.6458	0.6472
ETH-USD	5 min	Arbitrage spread	0.6126	0.4536	0.3806	0.3393	0.3158
ETH-USD	1 H	Order flow	0.1442	0.0844	0.0358	0.0433	**0.0032**
ETH-USD	1 H	Log-return	0.2039	0.2485	0.2372	0.2359	0.2334
ETH-USD	1 H	Volatility	0.9456	0.9373	0.9338	0.9292	0.9267
ETH-USD	1 H	Arbitrage spread	0.5443	0.3800	0.3723	0.2497	0.1892
ETH-USD	1 D	Order flow	**0.0751**	**− 0.0124**	**− 0.0513**	**0.0257**	**0.0618**
ETH-USD	1 D	Log-return	0.6550	0.5489	0.5756	0.6851	0.5071
ETH-USD	1 D	Volatility	0.9619	0.9055	0.8221	0.7359	0.6381
ETH-USD	1 D	Arbitrage spread	0.5414	0.1997	0.2292	0.2290	0.1689
BTC-USDT	5 min	Order flow	0.1959	0.1159	0.0947	0.0603	0.0497
BTC-USDT	5 min	Log-return	− 0.0210	− 0.0084	0.0115	− 0.0069	− 0.0166
BTC-USDT	5 min	Volatility	0.6334	0.4367	0.3632	0.3047	0.2725
BTC-USDT	5 min	Arbitrage spread	0.8490	0.7320	0.6623	0.6253	0.5829
BTC-USDT	1 H	Order flow	0.2044	0.1001	0.0858	0.0778	0.0575
BTC-USDT	1 H	Log-return	− 0.0340	− 0.0303	**0.0080**	0.0296	**− 0.0159**
BTC-USDT	1 H	Volatility	0.3986	0.2634	0.4046	0.1818	0.1584
BTC-USDT	1 H	Arbitrage spread	0.6125	0.3290	0.2500	0.2095	0.1887
BTC-USDT	1 D	Order flow	0.3312	0.3108	0.2613	0.3319	0.2531
BTC-USDT	1 D	Log-return	− 0.1407	**0.0755**	**− 0.0841**	**0.0750**	**− 0.0019**
BTC-USDT	1 D	Volatility	0.5364	0.1626	0.2472	0.2101	0.1133
BTC-USDT	1 D	Arbitrage spread	0.6282	0.4136	0.3642	0.3458	0.2808
ETH-USDT	5 min	Order flow	0.2111	0.1321	0.1087	0.0743	0.0641
ETH-USDT	5 min	Log-return	− 0.0143	− 0.0167	0.0215	− 0.0200	− 0.0235
ETH-USDT	5 min	Volatility	0.5709	0.3908	0.3217	0.2019	0.1962
ETH-USDT	5 min	Arbitrage spread	0.8811	0.7581	0.6694	0.5972	0.5280
ETH-USDT	1 H	Order flow	0.2561	0.1697	0.1693	0.1346	0.1001
ETH-USDT	1 H	Log-return	−0.0250	**− 0.0110**	**0.0121**	**0.0140**	**− 0.0209**

Table 11 (continued)

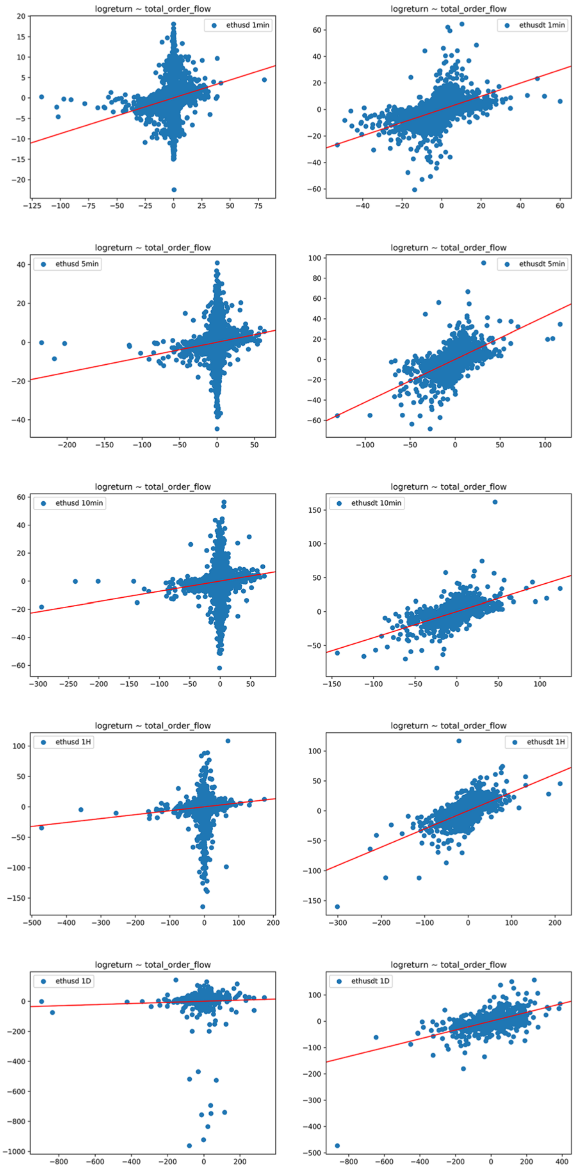
Symbol	Interval	Measure	t – 1	t – 2	t – 3	t – 4	t – 5
ETH-USDT	1 H	Volatility	0.3691	0.2067	0.3587	0.1774	0.1367
ETH-USDT	1 H	Arbitrage spread	0.5525	0.2211	0.1757	0.1542	0.1391
ETH-USDT	1 D	Order flow	0.4416	0.4277	0.4052	0.4576	0.3647
ETH-USDT	1 D	Log-return	−0.1301	**0.1014**	**− 0.0728**	0.1286	**0.0111**
ETH-USDT	1D	Volatility	0.5172	**0.0957**	0.1715	0.1776	**0.0763**
ETH-USDT	1 D	Arbitrage spread	0.4828	0.2800	0.3067	0.2856	0.2730
ETH-BTC	5 min	Order flow	0.1289	0.0778	0.0611	0.0628	0.0581
ETH-BTC	5 min	Log-return	**− 0.0035**	− 0.0195	− 0.0105	− 0.0102	**− 0.0008**
ETH-BTC	5 min	Volatility	0.4188	0.1753	0.1563	0.1063	0.1002
ETH-BTC	5 min	Arbitrage spread	0.7992	0.7044	0.6380	0.5927	0.5531
ETH-BTC	1 H	Order flow	0.3312	0.3129	0.2915	0.2784	0.2884
ETH-BTC	1 H	Log-return	− 0.0399	**− 0.0037**	**0.0063**	**− 0.0015**	**− 0.0106**
ETH-BTC	1 H	Volatility	0.3271	0.2242	0.2335	0.1536	0.1113
ETH-BTC	1 H	Arbitrage spread	0.7442	0.5781	0.4729	0.4030	0.3531
ETH-BTC	1 D	Order flow	0.8615	0.8074	0.7880	0.7616	0.7493
ETH-BTC	1 D	Log-return	**0.0859**	**− 0.0129**	**− 0.0091**	**− 0.0084**	**0.0111**
ETH-BTC	1 D	Volatility	0.4593	0.1147	0.1317	0.1376	**0.0968**
ETH-BTC	1 D	Arbitrage spread	0.6462	0.3429	0.3129	0.3010	0.2818
USDT-USD	5 min	Order flow	0.3451	0.1774	0.0789	0.0751	0.0621
USDT-USD	5 min	Log-return	**− 0.0071**	**− 0.0004**	0.0205	− 0.0189	**− 0.0030**
USDT-USD	5 min	Volatility	0.4856	0.4696	0.4318	0.4620	0.4477
USDT-USD	5 min	Arbitrage spread	0.2276	0.1638	0.1739	0.1188	0.1280
USDT-USD	1 H	Order flow	0.2327	0.1115	0.0868	0.0852	0.0811
USDT-USD	1 H	Log-return	**− 0.0010**	**− 0.0009**	− 0.0242	**0.0094**	0.0972
USDT-USD	1 H	Volatility	0.8044	0.7269	0.7073	0.6911	0.6763
USDT-USD	1 H	Arbitrage spread	0.3277	0.1726	0.3041	0.1948	0.0510
USDT-USD	1 D	Order flow	0.3330	0.2569	0.1227	**0.0908**	**− 0.0180**
USDT-USD	1 D	Log-return	**− 0.0873**	0.1393	− 0.1227	− 0.2293	**− 0.0548**
USDT-USD	1 D	Volatility	0.8003	0.7435	0.7524	0.7334	0.5215
USDT-USD	1 D	Arbitrage spread	0.4410	0.1626	0.1897	0.1241	**0.0467**

We consider up to 5 lags. Values in bold are not significant at the 1% level

Appendix 4. Scatter plots

See Figs. 6, 7, 8 and 9.

Fig. 6 Scatter plot of ETH-USD (left) and ETH-USDT (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: OF, *y*-axis: log-return. The straight line obtained from the linear regression is reported



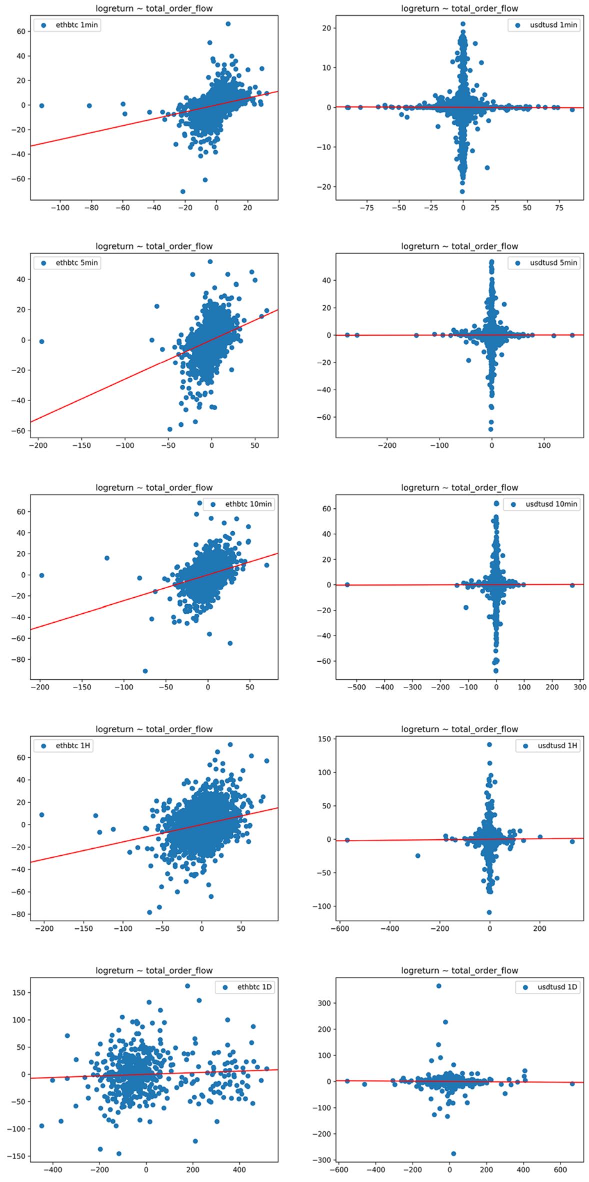


Fig. 7 Scatter plot of ETH-BTC (left) and USDT-USD (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: OF, *y*-axis: log-return. The straight line obtained from the linear regression is reported

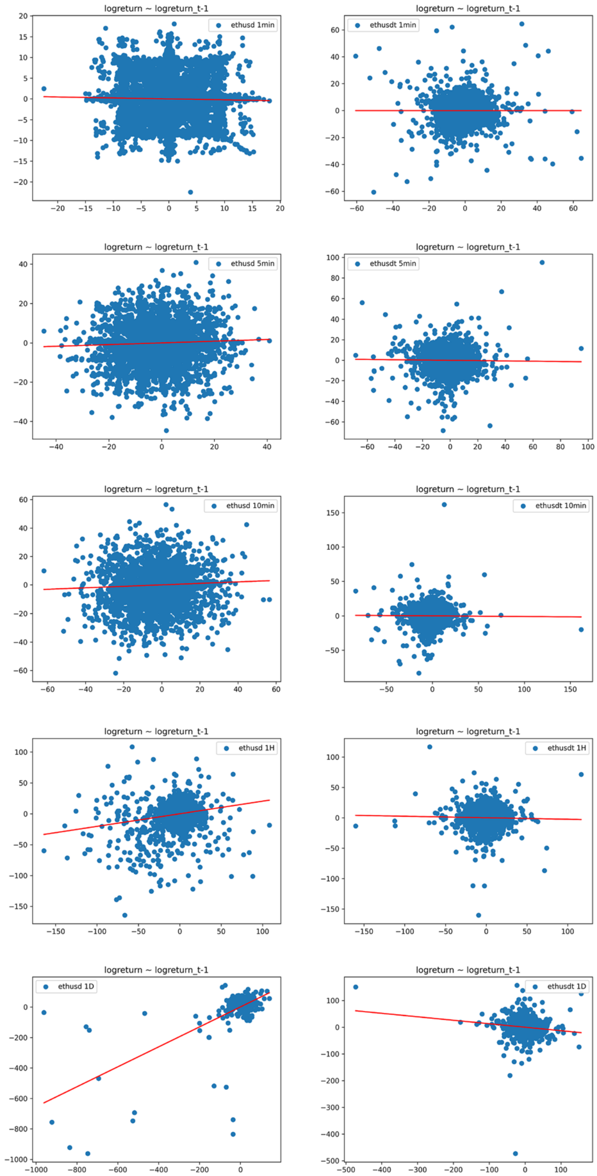


Fig. 8 Scatter plot of ETH-USD (left) and ETH-USDT (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: lagged log-return, *y*-axis: log-return. The straight line obtained from the linear regression is reported

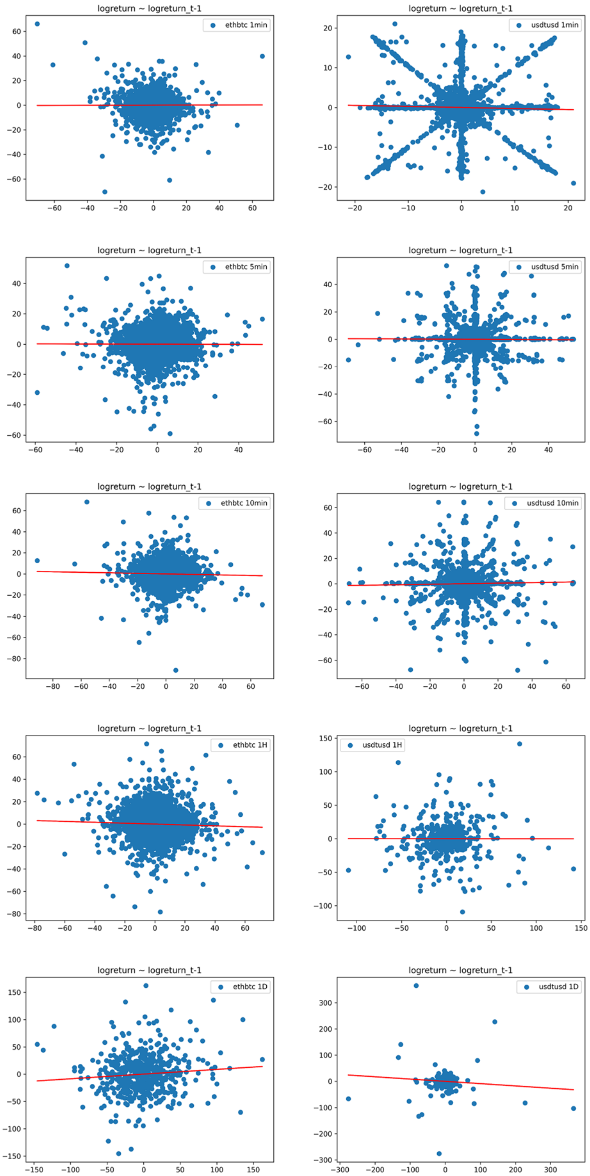


Fig. 9 Scatter plot of ETH-BTC (left) and USDT-USD (right): 1 min, 5 min, 10 min, 1 h, 1 day. *x*-axis: lagged log-return, *y*-axis: log-return. The straight line obtained from the linear regression is reported

Acknowledgements Emilio Barucci and Daniele Marazzina gratefully acknowledge support from the European Union's Horizon 2020 COST Action "FinAI: Fintech and Artificial Intelligence in Finance—Towards a transparent financial industry" (CA19130).

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The au-

thors have no competing interests to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alexander, C., & Dakos, M. (2020). A critical investigation of cryptocurrency data and analysis. *Quantitative Finance, 20*(2), 173–188.

Ballis, A., & Drakos, K. (2020). Testing for herding in the cryptocurrency market. *Finance Research Letters, 33*, 101210.

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

Barucci, E., Giuffra Moncayo, G., & Marazzina, D. (2022). Cryptocurrencies and stable coins: A high frequency analysis. *Digital Finance, 4*(2–3), 217–239.

Baur, D. G., Hong, K., & Lee, A. D. (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money, 54*, 177–189.

Berger, D. W., Chaboud, A. P., Chernenko, S. V., Howorka, E., & Wright, J. H. (2008). Order flow and exchange rate dynamics in electronic brokerage system data. *Journal of International Economics, 75*(1), 93–109.

Bouri, E., Gupta, R., & Roubaud, D. (2019). Herding behaviour in cryptocurrencies. *Finance Research Letters, 29*, 216–221.

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.

Brauneis, A., & Mestel, R. (2018). Price discovery of cryptocurrencies: Bitcoin and beyond. *Economics Letters, 165*, 58–61.

Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics and Data Analysis Nonlinear Modelling and Financial Econometrics, 51*(4), 2232–2245.

Chordia, T., & Subrahmanyam, A. (2004). Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics, 72*(3), 485–518.

Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. *Journal of Financial Economics, 65*(1), 111–130.

Chordia, T., Roll, R., & Subrahmanyam, A. (2005). Evidence on the speed of convergence to market efficiency. *Journal of Financial Economics, 76*(2), 271–292.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2018). Virtual relationships: Short-and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money, 52*, 173–195.

Cont, R., Kukanov, A., & Stoikov, S. (2013). The price impact of order book events. *Journal of Financial Econometrics, 12*(1), 47–88.

De Jong, A., Rosenthal, L., & Van Dijk, M. A. (2009). The risk and return of arbitrage in dual-listed companies. *Review of Finance, 13*(3), 495–520.

Donier, J., & Bonart, J. (2015). A million metaorder analysis of market impact on the Bitcoin. *Market Microstructure and Liquidity, 01*(02), 1550008.

Evans, M. D. D., & Lyons, R. K. (2002). Order flow and exchange rate dynamics. *Journal of Political Economy, 110*(1), 170–180.

Ghadhab, I., & Hellara, S. (2015). The law of one price, arbitrage opportunities and price convergence: Evidence from cross-listed stocks. *Journal of Multinational Financial Management, 31*, 126–145.
Goyenko, R. Y., Holden, C. W., & Trzcinka, C. A. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*(2), 153–181.
Griffin, J., & Shams, A. (2020). Is Bitcoin really untethered? *The Journal of Finance, 75*(4), 1913–1964.
Grobys, K. (2021). When the blockchain does not block: On hackings and uncertainty in the cryptocurrency market. *Quantitative Finance, 21*(8), 1267–1279.
King, T., & Koutmos, D. (2021). Herding and feedback trading in cryptocurrency markets. *Annals of Operations Research, 300*, 79–96.
Lintilhac, P. S., & Tourin, A. (2017). Model-based pairs trading in the Bitcoin markets. *Quantitative Finance, 17*(5), 703–716.
Lyons, R. K., & Viswanath-Natraj, D (2019). What keeps stable coins stable? SSRN ID 3508006. *Social Science Research Network*.
Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics, 135*(2), 293–319.
Manahov, V. (2021). Cryptocurrency liquidity during extreme price movements: Is there a problem with virtual money? *Quantitative Finance, 21*(2), 341–360.
Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2013). ETF arbitrage: Intraday evidence. *Journal of Banking and Finance, 37*(9), 3486–3498.
Nadarajah, S., & Chu, J. (2017). On the inefficiency of Bitcoin. *Economics Letters, 150*, 6–9.
Petukhina, A., Trimborn, S., Härdle, W. K., & Elendner, H. (2021). Investing with cryptocurrencies-evaluating their potential for portfolio allocation strategies. *Quantitative Finance, 21*(11), 1825–1853.
Roll, R., Schwartz, E., & Subrahmanyam, A. (2007). Liquidity and the law of one price: The case of the Futures-Cash basis. *The Journal of Finance, 62*(5), 2201–2234.
Sensoy, A. (2019). The inefficiency of Bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters, 28*, 68–73.
Silantyev, E. (2019). Order flow analysis of cryptocurrency markets. *Digital Finance, 1*(1), 191–218.
Tiwari, A. K., Jana, R., Das, D., & Roubaud, D. (2018). Informational efficiency of Bitcoin—An extension. *Economics Letters, 163*, 106–109.
Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters, 148*, 80–82.
Vidal-Tomás, D. (2021). An investigation of cryptocurrency data: The market that never sleeps. *Quantitative Finance, 21*(12), 2007–2024.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.