The Dual Impact of On-chain and Off-chain Factors on Bitcoin Market Efficiency

Weiwei Guo[a], Hossein Jahanshahloo[b,*], Laima Spokeviciute[a], Qingwei Wang[a]

[a] *Cardiff Business School, Cardiff University, Cardiff, UK.*
[b] *Alliance Manchester Business School, The University of Manchester, Manchester, UK.*

Abstract

This paper examines how on-chain factors (number of active wallets, transaction fees, and transaction volume) and off-chain factors (liquidity and investor attention) impact Bitcoin market efficiency from April 2014 to April 2022. We identify three periods in Bitcoin's market development: development, growth, and additional development stage. We propose three hypotheses: (1) increased investor attention enhances market efficiency, (2) a rise in active users improves efficiency directly and through liquidity and investor attention, and (3) higher transaction fees and on-chain volume positively impact efficiency directly and indirectly. Our findings support these hypotheses during Bitcoin's development and growth periods. However, in the additional development stage, the total effect of active users, transaction fees, and transaction volume becomes negative when considering mediating effects, and largely insignificant when focusing on direct effects. Additionally, we find increased netflow between whales and exchanges, a proxy for institutional activity, improves efficiency. We conclude that as Bitcoin's market develops, factors such as changing user composition and increased regulatory scrutiny alter the dynamics of on-chain factors and their influence on market efficiency.

Keywords: Bitcoin, Market efficiency, Blockchain, Random Walk.

JEL Codes: G12, G14

[⋆] We thank conference participants at Edinburgh Economics of Financial Technology Conference 2023, Cardiff Fintech Conference 2023, Welsh Postgraduate Research Conference 2023, Cardiff Business School Accounting and Finance Seminar, 2024 Lancaster-Manchester-Warwick Joint PhD Workshop, 2024 Edinburgh World-Class Workshop, Tri-University (Cardiff-Xiamen-Newcastle) Annual Conference 2024, and FEBS MENA-Asian 2024 Conference. All remaining errors are our own.

[*] Corresponding author

Email addresses: `GuoW17@cardiff.ac.uk` (Weiwei Guo), `Hossein.Jahanshahloo@manchester.ac.uk` (Hossein Jahanshahloo), `SpokeviciuteL@cardiff.ac.uk` (Laima Spokeviciute), `WangQ30@cardiff.ac.uk` (Qingwei Wang)

1. Introduction

According to Fama (1970), the efficient market hypothesis (EMH) suggests that in an efficient market, the returns are unpredictable since prices have already reflected all the available information. With the growing popularity of cryptocurrencies, especially Bitcoin, academics have also sought to understand whether and how the EMH works in the Bitcoin market. Urquhart (2016) is among the first to study EMH in the Bitcoin market and finds that Bitcoin market efficiency improves as it develops. Later studies have further shown that the Bitcoin market efficiency is indeed time-varying. For example, Khuntia and Pattanayak (2018) find that Bitcoin returns exhibit time-varying linear and nonlinear dependence. More recently, Kakinaka and Umeno (2022) find that after the outbreak of COVID-19, Bitcoin market efficiency level decreases in the short-time horizon, and increases in the long-time horizon.

A separate strand of literature further investigates the determinants of Bitcoin market efficiency and the main focus is drawn on market-based factors such as liquidity and volatility. Specifically, these studies show that decreased Bitcoin return predictability, as an indicator of improved market efficiency, has been associated with increased liquidity (see, for example, Brauneis and Mestel (2018), Wei (2018), and Al-Yahyaee et al. (2020)) and reduced volatility (see, for example, Al-Yahyaee et al. (2020)). These findings are consistent with Chordia et al. (2008) and Rösch et al. (2017)), who study traditional financial markets. Chordia et al. (2008) show that short-term return predictability of NYSE firms reduces and prices become closer to a random walk process as the bid-ask spread narrows, and Rösch et al. (2017) find that increased volatility leads to increased return predictability in the U.S. stock market.

Another factor that has been widely discussed in traditional financial markets is investor attention. Vozlyublennaia (2014) have documented that increased retail investor attention reduces return predictability, leading to higher efficiencies in the stock, bond, gold, and oil markets. Previous studies on Bitcoin investor attention mainly focus on its relationships with returns, trading volume, volatility, and liquidity (Urquhart, 2018; Shen et al., 2019; Liu and Tsyvinski, 2021; Smales, 2022), but not with market efficiency. Urquhart (2018) and Shen et al. (2019) find daily investor attention does not predict future returns, whereas Liu and Tsyvinski (2021) find it has strong predictability for returns in weekly frequency. Urquhart (2018) and Shen et al. (2019) show bidirectional predictive relationships between investor attention and both trading volume and volatility, and Smales (2022) finds that increased investor attention is associated with increased illiquidity and volatility. However, in Smales (2022), illiquidity is estimated with Amihud (2002) illiquidity ratio, which performs poorly in capturing the time-series variability of Bitcoin market liquidity (Brauneis et al., 2021). These complex interactions can also be found in traditional financial markets. For example, while investor attention is positively related to volatility (Vlastakis and Markellos, 2012), potentially undermining market efficiency, it nonetheless contributes to enhanced market efficiency in the U.S. stock market (Vozlyublennaia, 2014).

Theoretically, investor attention can influence market efficiency through two contrasting mechanisms, one enhancing efficiency and the other diminishing it. According to Grossman and Stiglitz (1980), prices reflect information more accurately, which implies a higher price efficiency, when more investors are informed. Increased investor attention enhances information gathering and accelerates its incorporation into prices, reducing return predictability and improving market efficiency. Grossman and Stiglitz (1980) theory is supported by Vozlyublennaia (2014) empirical study in stock, bond, gold, and oil markets. On the other hand, Da et al. (2011) suggest that increased retail investor attention induces more noise into price, which decreases market efficiency. This aligns with Barber and Odean (2008) theorem, which states that retail investors are more likely to buy than sell securities that capture their attention. Consequently, increased investor attention drives buying pressure, leading to short-term momentum and long-term price reversals. Similarly, in the cryptocurrency market, Sockin and Xiong (2023) model suggests investor attention can drive momentum through incorrect expectations about future prices. However, empirically, Liu and Tsyvinski (2021) find that investor attention on Bitcoin has no significant impact on time-series momentum. This is consistent with Kogan et al. (2023) finding that inattention does not explain Bitcoin traders' momentum strategy. Building on investor attention's impact on market efficiency in traditional financial markets, and its relationships with Bitcoin market activity and dynamics, we construct our first hypothesis.

Hypothesis 1: Investor attention has a positive effect on Bitcoin market efficiency.

This hypothesis is further motivated by factors beyond the established associations. For example, despite the growing popularity of Bitcoin, its fundamental value remains an open question. Pagnotta (2022) believes that Bitcoin's fundamental value is rooted in its decentralized security system. This system involves a balance of user transactions, miner activities, and the risk of attacks, resulting in varying price levels. Biais et al. (2023) model suggests that Bitcoin's fundamental value is its net transactional benefits. The unresolved fundamental value makes Bitcoin pricing more subjective, which emphasizes investor attention's importance.

Although Bitcoin shares similar features with traditional financial assets, it is built on fundamentally different technology, blockchain. These blockchain-based factors play an important role in Bitcoin market. For example, Liu and Tsyvinski (2021) find that blockchain-based factors have strong correlations with and can be used to predict Bitcoin returns. Furthermore, blockchain-based factors have significant connections with market-based factors that in turn affect market efficiency. For example, Brauneis et al. (2022) find that increased transaction fees and on-chain transaction volume are associated with improved liquidity. As a result of these studies, we employ in our analysis of Bitcoin market efficiency blockchain-based factors, including the number of active wallets, transaction fees, and on-chain transaction volume.

In Bitcoin, users can and have multiple addresses in their wallets. Using the Union-Find Algorithm, Foley et al. (2019) aggregate addresses into wallets, and use the number of wallets as a proxy for the number of users to pinpoint Bitcoin involved illegal activities. A later study by Jahanshahloo et al. (2023) shows that wallets are indeed a better representation of active users than addresses because Bitcoin blockchain users frequently transfer their Bitcoin to new addresses to increase security and privacy. Therefore, we use the number of active wallets to estimate the number of active users in Bitcoin blockchain. In the stock market, Eaton et al. (2022) find that decreased participation of Robinhood investors is associated with increased liquidity and decreased volatility for stocks that attract retail investors. Bitcoin is dominated by retail investors (Jahanshahloo et al., 2023) and Bitcoin investors exhibit similar stock holding patterns to Robinhood investors (Kogan et al., 2023). Given the aforementioned evidence on liquidity and volatility effects on efficiency, we hypothesize that there is a direct connection between the number of active wallets and Bitcoin market efficiency, as well as an indirect connection that is mediated by liquidity and investor attention:

Hypothesis 2a: An increase in the number of active wallets improves Bitcoin market efficiency.

Hypothesis 2b: An increase in liquidity and investor attention, and a decrease in volatility positively mediates the relationship between the number of active wallets and Bitcoin market efficiency.

Transaction fees in the Bitcoin market are the payments made to the Bitcoin miners to verify and include a transaction in a block. While these fees are typically viewed as a cost that could either increase or decrease market efficiency in different markets[1], they are also indicators of congestion in Bitcoin blockchain because higher fees are paid in order to avoid prolonged delays in the processing of the transactions (Easley et al., 2019; Huberman et al., 2021). Therefore, higher fees usually are associated with increased on-chain transaction activities, which in turn is associated with improved liquidity Brauneis et al. (2022). Following these studies, we conjecture that:

Hypothesis 3a: An increase in the transaction fees have a positive effect on Bitcoin market efficiency.

Hypothesis 3b: An increase in liquidity positively mediates the relationship between transaction fees and Bitcoin market efficiency.

In terms of on-chain transaction volume, we expect it to have a positive effect on Bitcoin market efficiency. We base our opinion on the findings by Brauneis et al. (2022) and Aalborg et al. (2019). Specifically, Brauneis et al. (2022) study the period from 12/15/2017 to 12/15/2019 and find that on-chain transaction volume is positively related to Bitcoin liquidity. Meanwhile, Aalborg et al.

[1] From the cost perspective, transaction fees are similar to the financial transaction tax (FTT) in stock market. Colliard and Hoffmann (2017) find FTT decreases market efficiency in French market. However, in the US market, Cipriani et al. (2022) find that FTT improves market efficiency.

(2019) demonstrate that Bitcoin on-chain transaction volume has a significant positive correlation with Bitcoin trading volume, and the latter is positively related to investor attention (Shen et al., 2019). Therefore, our next set of hypotheses state the following:

Hypothesis 4a: An increase in on-chain transaction volume has a positive effect on Bitcoin market efficiency.

Hypothesis 4b: An increase in on-chain transaction volume has a positive association with market efficiency because of its relationship with liquidity and investor attention, introduced by trading volume.

We proceed to test our hypotheses in the following steps. In line with the findings on time-variation of Bitcoin price efficiency, we first show that for our sample period between April 10, 2014 and April 30, 2022, there are two structural breaks in the time series of Bitcoin market efficiency. The first period (between April 10, 2014 and 14 July 2017) can be categorized as the development stage, with low market efficiency and liquidity, low fees and transaction volume. In this period, Bitcoin was largely seen as a tool for anonymous and decentralized transactions and primarily utilized as a medium of exchange for illegal activities (Foley et al., 2019). The second period (between 14 July 2017 and 16 November 2020) is when Bitcoin was at its growth stage, with lower, however, more volatile market efficiency, higher liquidity, fees and transaction volume. During this period, despite a significant decline in the proportion of illegal use, Bitcoin remained far from a traditional financial asset, as its market was still largely driven by speculative interest (Foley et al., 2019; Makarov and Schoar, 2021). In the third period (between 17 November 2020 and April 30, 2022) Bitcoin market reached an additional development stage, and its efficiency increased, together with liquidity and the on-chain factors. By this time, Bitcoin was increasingly recognized as a financial asset, driven by the rising adoption among institutional investors (Bianchi and Babiak, 2022; Huang et al., 2022; Momtaz, 2024) and regulatory scrutiny (Adrian et al., 2022; Iyer, 2022).

To test the direct effects between the on-chain and off-chain factors and Bitcoin market efficiency, we run a set of regressions, where as in Comerton-Forde and Putniņš (2015) we employ Lo and MacKinlay (1988) variance ratio test and return autocorrelation as the dependent variables to measure Bitcoin market efficiency. To account for the structural breaks, we first estimate the results for the full sample, including dummy variables that represent the distinct periods, and later run the regressions for the three sub-samples individually. Due to high correlations, we report the results for each of the on-chain factors separately.

Our findings indicate that the hypothesized relationships hold differently across sub-samples. More specifically, we find that increased investor attention is associated with increased Bitcoin market efficiency in the first two periods which is consistent with our Hypothesis H1 and Vozlyublennaia (2014) findings in various traditional financial markets, and supports the view that increased investor attention brings more information to the market. However, in the final period the positive impact of investor attention on market efficiency diminishes due to the Bitcoin bullish phase, influenced by the COVID-19 pandemic and heightened market hype (Biktimirov and Biktimirova, 2023). During this time, sentiment played a key role in driving prices (Long et al., 2024; Osman et al., 2024). While investor attention typically enhances information flow, it can also introduce noise (Da et al., 2011). In this period, sentiment-fueled speculation likely amplified noise, offsetting the benefits of information flow and neutralizing the effect of investor attention.

The increased number of active wallets is associated with increased market efficiency for the first period, while shows a negative effect on market efficiency for the later periods. The shift in their relationship is consistent with the change in Bitcoin user composition. In the early stage, 46% of Bitcoin transactions were by illegal users using it as a mean of payment (Foley et al., 2019), which makes them intrinsic users because the payment system plays a significant role in Bitcoin's fundamental value (Pagnotta, 2022; Biais et al., 2023). However, as the interests of other market participants grew, the proportion of illegal users dropped, and by 2017 the majority of Bitcoin users were formed by speculators and investors (Makarov and Schoar, 2021), and the market became dominated by retail traders (Jahanshahloo et al., 2023; Kogan et al., 2023). This explains the negative relation between the number of active wallets and market efficiency since increased number of active users brings excess noise.

Similarly, this user composition change alters Bitcoin market efficiency relationships with transaction fees and on-chain transaction volume as their increases are associated with improved market efficiency only when Bitcoin was in the development stage. In the later periods, this relationship becomes insignificant, suggesting other factors, such as growing Bitcoin trading activities via exchanges[2] (i.e. off-chain trading, where the transaction do not need to be recorded on the blockchain) reduced their impact. Overall, our direct effect hypotheses H2a, H3a and H4a for on-chain factors are only supported for period 1.

We next proceed with the mediation analysis to test the channels through which the on-chain factors affect Bitcoin market efficiency. Again, we observe different results for the three distinct Bitcoin market development periods. In the first two periods, the overall effects of on chain factors (direct and indirect ones) are positive, supporting our indirect effect hypotheses, H2B, H3b and H4b. In other words, the on-chain factors have a positive effect on liquidity and investors' attention, which in turn increases Bitcoin market efficiency.

In the final period, however, the total impact is negative and significant, and this effect is mainly dominated by the liquidity channel. The negative relationship between liquidity and transaction activity is due to "fear of missing out" of retail investors Baur and Dimpfl (2018), which leads to heightened trading demand in Bitcoin bull cycle. This surge in trading activity results in a sharp drop in liquidity, resembling patterns observed in the stock market Yeyati et al. (2008).

[2]For the increase of role of exchanges, see Jahanshahloo et al. (2023).

Our study contributes to the study of the interaction of Bitcoin on-chain and off-chain characteristics, and specifically helps understand the evolution of Bitcoin market efficiency. We examine factors affecting Bitcoin market efficiency not only from the market perspective but also from Bitcoin blockchain perspective, and further investigate blockchain-based factors' relationship with market efficiency when considering their interaction with market-based factors. This sheds light on the dynamics of Bitcoin market, as an asset that shares similar features with that of a stock but is built on fundamentally different technology, blockchain. We find that both market-based and blockchain-based factors have significant relationships with Bitcoin market efficiency. This contributes to the debate on what drives Bitcoin market efficiency, which is crucial for investors determining their trading strategies as well as regulators trying to improve market quality.

The rest of the paper is structured as follows. In the next section we present our data sample and methodology. Section 3 presents our results and discussion, and the final section concludes.

2. Data Sample and Methodology

2.1. Data Sample

Following Jahanshahloo et al. (2023), we collect historical Bitcoin prices in U.S. dollar on available exchanges from BitcoinCharts.com and calculate the volume-weighted-average price (VWAP) across all exchanges at each minute to represent the trade price. To estimate investor attention, we follow Shen et al. (2019) and use daily number of tweets on "Bitcoin" from BitInfoCharts, which covers data from April 2014 onwards. We obtain on-chain data, including number of active wallets, transaction fees, transaction volume, and transaction weights in each block, and the netflow for whales (calculated as flow from exchange to whales minus flow from whales to exchange) from Cardiff University Bitcoin Database[3] (CUBiD), which covers January 3, 2009 - April 30, 2022. After merging all the datasets, our final sample period spans from April 10, 2014, to April 30, 2022.

2.2. Methodology

We base our hypotheses on the premise that bitcoin market efficiency is either directly or indirectly affected by on-chain factors. These theoretical paths of effect are illustrated in Figure 1 below. The direct effects of off-chain and on-chain factors on market efficiency are depicted by arrows a $\rightarrow$ b, and c, respectively. To test these paths, and the related Hypotheses H1, H2a, H3a and H4a, we set up the following baseline OLS regression:

$$Market\ Efficiency_t = \beta_{10} + \beta_{11} Off\text{-}chain_t + \beta_{12} On\text{-}chain_t + WhalesNetflow_t + \epsilon_1 \tag{1}$$

[3]https://www.cardiff.ac.uk/business-school/collaborate/cubid

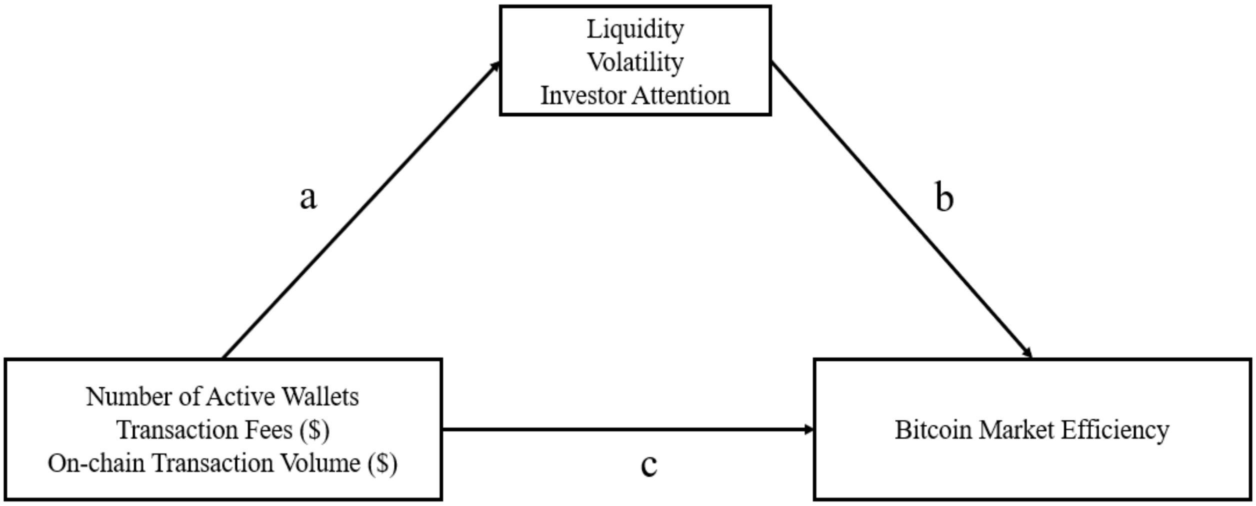

Figure 1: Causal chains of Bitcoin market efficiency, on-chain and off-chain factors in the Mediation Model
This figure shows the theoretical path of our methodology for estimating the impact of on-chain factors on off-chain factor and market efficiency and their mediation through off-chain factors.

where *Market Efficiency*$_t$ is one of our estimate of market efficiency on day t, and *Off-chain*$_t$ and *On-chain*$_t$ are the vectors for our off-chain and on-chain measures, respectively, $WhalesNetflow_t$ is the netflow between whales and exchanges, and ϵ_1 is the standard robust error term.

The indirect effects of on-chain factors on market efficiency are depicted by the pathway a through b, i.e. on-chain factors affect bitcoin market efficiency via their impact on off-chain factors. A formal mediation test includes running the following set of regressions:

$$\textit{On-chain}_t = \beta_{20} + \beta_{21}\textit{Off-chain}_t + WhalesNetflow_t + \epsilon_2 \tag{2}$$

where *On-chain*$_t$ is one of the on-chain factors.

We run the mediation analysis using the structural equation modelling (SEM) method, which fits direct and indirect paths simultaneously (Iacobucci et al., 2007). To estimate the indirect effect and test its significance, we employ delta method (Cramér, 1999), Sobel test (Sobel, 1982), and Monte Carlo simulation with 10,000 replications. Finally we estimate the proportional contributions of both direct and indirect effects to bitcoin market efficiency.

2.3. Market Efficiency

Following Chordia et al. (2008) and Comerton-Forde and Putniņš (2015), we employ variance ratio (Lo and MacKinlay, 1988) and return autocorrelation as proxies for market efficiency. For a random walk process, the variance of kl-period returns is k times the variance of l-period returns. Based on this, a larger deviation of the variance ratio from the random walk benchmark, which is one, indicates a less efficient market (Lo and MacKinlay, 1988). The formal expression for our daily market efficiency using variance ratio (day subscripts are suppressed) is as follows:

$$VR_{kl} = -|\frac{\sigma_{kl}^2}{k\sigma_l^2} - 1| \tag{3}$$

where σ_l^2 and σ_{kl}^2 are the daily variances of l-minute and kl-minute returns, and k is the quotient of kl and l.

To make it more intuitive, our market efficiency measure takes the absolute value of the difference between variance ratio test statistics and 1, and then multiply it by -1. Therefore, a higher value of VR_{kl} indicates the prices are closer to the random walk benchmark, which means that market efficiency is higher. Following Comerton-Forde and Putniņš (2015), we use three different (l, kl) combinations, (1-minute and 5-minute), (5-minute, 30-minute), and (30-minute, 60-minute). We compute the market efficiency measure, VR, by taking the first principal component of the three variance ratios, and then scaling it to range from zero (highly inefficient) to one (highly efficient).

For autocorrelation, both positive and negative values indicate an inefficient market, as returns should have no autocorrelation if they are not predictable using past information (Chordia et al., 2008; Comerton-Forde and Putniņš, 2015). Following Comerton-Forde and Putniņš (2015), we calculate the first-order autocorrelations of 1-min, 5-min, and 30-min return for each day as:

$$Autocorrelation_t = Corr(R_\tau, R_{\tau-1}) \tag{4}$$

Then, we take the opposite of the first principal component of their absolute values, and scale it to range from zero (highly inefficient) to one (highly efficient) to obtain the market efficiency measure, AC.

Bitcoin 1-, 5-, 30-, and 60-minute returns are calculated as:

$$R_\tau = ln(\frac{P_\tau}{P_{\tau-1}}) \times 100 \tag{5}$$

where P_τ is Bitcoin price for each intraday frequency.

To remove potential errors in the high-frequency Bitcoin price, we clean the price data following the procedures in Brownlees and Gallo (2006) before performing further analyses. Specifically, we replace the price with the price at the last minute if the absolute difference between it and the average price within a nearby 30-minute neighborhood is larger than 3 times the standard deviation. However, different from Brownlees and Gallo (2006), we do not limit the neighborhood of observations on the same day because Bitcoin is traded 24 hours a day and seven days a week.

2.4. Off-chain Factors

To estimate investor attention, we use the total number of tweets posted on Bitcoin (#Tweets*). To measure liquidity, follwoing Scharnowski and Jahanshahloo (2025) we use the Corwin and Schultz (2012) bid-ask spread estimator (CS), because it captures the time-series variability of liquidity in the cryptocurrency market better than other liquidity estimators that are commonly used in the stock market, such as the Amihud illiquidity ratio (Brauneis et al., 2021). Specifically, the liquidity for day t is estimated using the following set of equations:

$$\begin{aligned}
Liquidity_t &= -\frac{1}{n-1}\sum_{1}^{n-1} max(CS_{i,i+1}, 0) \\
CS_{i,i+1} &= \frac{2(\exp(\alpha)-1)}{1+\exp(\alpha)} \\
\alpha &= \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}} \\
\beta &= [ln(\frac{H_i}{L_i})]^2 + [ln(\frac{H_{i+1}}{L_{i+1}})]^2 \\
\gamma &= [ln(\frac{H_{i,i+1}}{L_{i,i+1}})]^2
\end{aligned} \tag{6}$$

where H_i and L_i are respectively the highest and lowest prices in every one-hour sub-interval, i, for every day, t.

Following Corwin and Schultz (2012) and Brauneis et al. (2021), we set the negative values of $CS_{i,i+1}$ to zero and liquidity at day t is the opposite of the unweighted average of $CS_{i,i+1}$ within day t. We multiply the liquidity estimate by -1 for more intuitive interpretation, i.e. higher values now indicate that the market is more liquid. Same with Brauneis et al. (2021), we do not adjust the estimator for the overnight trading halt because Bitcoin is traded 24 hours a day and seven days a week.

Following Urquhart (2018), we measure volatility with realized volatility as follows:

$$RV_t = \sqrt{\sum_{i=1}^{N} R_{t,i}^2} \tag{7}$$

where $R_{t,i}$ is the 1-minute log return of Bitcoin in sub-interval i at day t, and N is the number of intraday return intervals

2.5. On-chain Factors

CUBiD provides the number of active wallets, transaction volume, and transaction fees at the block level. Different wallets are identified via the Union-Find Algorithm (UFA) which combines addresses belonging to the same user[4]. Therefore, we aggregate the number of active wallets in each block within one day to estimate the daily number of active users in thousands (#Active Wallets*. As Bitcoin price changes drastically, analyzing its transaction volume and fees in fiat currency, rather than in Bitcoin units, aligns more closely with investors' concerns by accounting for the real-world cost implications that vary widely with Bitcoin's market value. Therefore, following Jahanshahloo et al. (2023), we use the last VWAP before or at the mining time of a block to calculate the transaction volume and transaction fees in USD in each block. We then take the sum of transaction volume in USD for all the blocks in a day as the daily Transaction Volume*[5] and present it in billions of USD. We calculate the daily transaction fees per weight (FpW*) as the ratio of the sum of the total transaction fees and the sum of the transaction weights in a day. This method captures the detailed variability of fees across all daily transactions.

The daily whales netflow is calculated by identifying wallets holding over 5,000 BTC at any point during the day and excluding the known entities such as exchanges, services, gambling websites, mixers, etc. For these wallets, the netflow is the difference between the Bitcoin received from and sent to identified exchanges, and we present it in billions of USD. Although Augustin et al. (2023) define wallets with over 1,000 BTC as "whales" using this as a proxy for large investors, we adopt a threshold of 5,000 BTC to represent institutional investors. This threshold is chosen because a holding of 5,000 BTC signifies institutional-level capital, indicating long-term, strategic Bitcoin investments. Such large holders can significantly influence market prices through substantial trades affecting Bitcoin's liquidity and market dynamics [6].

2.6. Summary Statistics

Because the number of tweets as well as the on-chain factors may exhibit general upward trends as Bitcoin gains popularity and trading activity increases, we detrend these variables to ensure we eliminate spurious correlation. Specifically, we follow Griffin et al. (2007), and detrend them by first taking the natural logarithm, then subtracting the mean of the prior 100-day data.

Table 1 provides the summary statistics for all the raw variables and the measures as they enter the baseline regression model in Equation (1). The average VWAP Bitcoin price during our sample period was \$11,438 , with the minimum value of \$167 on the 7th November, 2014, reaching the maximum of \$68,908 on the 10th November, 2021. The average number of tweets on Bitcoin in a day was 48,151 throughout our sample period, which suggests an extremely high investor attention. For comparison, Chen and Hwang (2022) report that the number of Tweets in a whole year in 2018 on Apple Inc., which is one of the most popular tickers in their sample, was 276,256

[4]See Foley et al. (2019) for more details on UFA.

[5]For robustness we conduct our analysis with the average on-chain transaction volume and the results are identical.

[6]For robustness, we also conducted our analysis using the 1,000 BTC threshold, and the results were consistent with those obtained using the 5,000 BTC threshold. Please refer to Appendix I for detailed results.

that is almost 5 times less per day than number of tweets on Bitcoin. The number of active wallets, which we use as a proxy for active users ranged between over 104 thousand and over 1 million, with an average of approximately 452 thousand, which demonstrates the significant variation in number of active users on Bitcoin throughout the time. Finally, the average transaction fees per weight was $0.001 and the average transaction volume in a day was $8.33 billion. All the raw variables, apart from the number of active wallets, have much higher means than medians, with large standard deviations, suggesting right-skewed distribution with some extreme high values.

Figure 2 provides a time-series of our main variables, liquidity, number of tweets (#Tweets), number of active wallets (#Active Wallets), transaction fees per weight (FpW), transaction volume, and whales netflow from April 10, 2014 to April 30, 2022. The vertical dashed lines in each panel indicate structural breaks in market efficiency with details explained in Figure 3. Liquidity reflects the ease of trading in the Bitcoin market, and shows an overall improving trend over time, particularly after 2020. Number of tweets captures investor attention, with prominent spikes aligning with major Bitcoin events, such as the 2017 and 2021 bubbles. Number of active wallets indicates user engagement, which generally trends upward over time with surges during high-activity periods. Transaction fees per weight reflects blockchain network demand, particularly peaking during periods of congestion in 2017 and 2021. Transaction volume captures daily Bitcoin transfer activity within the blockchain, with notable increases in recent years as adoption and usage have grown. Finally, institutional investor transaction activity, measured as whales' netflow, shows a significant surge immediately after the second structural break at the end of 2020.

We report the correlation matrix in Table 2. It shows that market efficiency is positively (negatively) related to liquidity (volatility), number of tweets, transaction volume, and whales netflow, which provides preliminary support for our first and last hypotheses. However, the negative correlations of the market efficiency with the number of active wallets and transaction fees contradict Hypotheses 2 and 3. In addition, the absolute correlation between liquidity and volatility is high. To avoid collinearity, we separate these variables into different regressions, focusing on liquidity in the main analysis. The results for volatility are robust and presented in Appendix II.

We plot the estimated Bitcoin market efficiency from April 10, 2014 to April 30, 2022 in Figure 3. We show that Bitcoin market efficiency varies over time, and more importantly, there seem to be structural breaks in the data over the sample period. We formally test for structural breaks with the Bai and Perron (1998) test, and find that the break dates are July 14, 2017, and November 16, 2020 (indicated with vertical lines in the figure).

Table 3 further reports the univariate tests for all variables used in the empirical analysis before detrending across the three identified distinct periods[7]. This a-priori test shows that the variables of interest are in fact different in the three periods. Specifically, we show that all variables have a significant upward trend throughout the time, with efficiency increasing by up to 10 fold, and

[7] We also performed Jonckheere–Terpstra (JT) test and the results are consistent.

Table 1: Summary Statistics
This table reports number of observations, means, medians, standard deviations, minimum and maximum values of both raw variables and measures used in the empirical analysis. Bitcoin price is in 1-minute frequency and other variables are in daily frequency. Bitcoin price and transaction fees per weight (FpW*) are in USD, number of active wallets (#Active Wallets) is in thousands, and Transaction Volume* and Whales Netflow are in billions of USD. VR and AC are market efficiency measures based on variance ratio test and autocorrelation respectively. The sample is from April 10, 2014 to April 30, 2022.

Raw Variables						
	N	Mean	p50	SD	Min	Max
Bitcoin Price ($)	4,237,920	11,438.00	5,788.00	16,327.66	167	68,908.00
#Tweets*	2,943	48,151.03	29,244.00	43,403.01	7,300.00	363,566.00
#Active Wallets*	2,943	451.98	476.18	165.38	104.37	1,001.89
FpW* ($)	2,943	0.001	0.0004	0.003	0.00001	0.04
Transaction Volume* ($)	2,943	8.33	3.40	17.46	0.08	380.70
Measures used in the Empirical Analysis						
	N	Mean	p50	SD	Min	Max
VR	2,943	0.26	0.11	0.29	0	1
AC	2,943	0.33	0.21	0.30	0	1
Liquidity	2,943	-0.02	-0.02	0.02	-0.14	-0.0006
Volatility	2,943	0.48	0.34	0.42	0.01	2.96
#Tweets	2,942	0.04	0.03	0.25	-1.11	1.48
#Active Wallets	2,942	0.00	0.00	0.01	-0.07	0.05
FpW	2,942	0.01	0.00	0.06	-0.21	0.25
Transaction Volume	2,942	0.01	0.01	0.05	-0.15	0.36
Whales Netflow	2,943	0.02	0.001	0.16	-1.89	4.17

Table 2: Correlation Matrix
This table reports the correlations between measures used in the empirical analysis. The sample is from April 10, 2014 to April 30, 2022.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) VR	1.00								
(2) AC	0.97	1.00							
(3) Liquidity	0.62	0.65	1.00						
(4) Volatility	-0.62	-0.66	-0.97	1.00					
(5) #Tweets	0.24	0.24	0.18	-0.17	1.00				
(6) #Active Wallets	-0.17	-0.17	-0.12	0.13	0.20	1.00			
(7) FpW	-0.07	-0.07	-0.07	0.06	0.35	0.37	1.00		
(8) Transaction Volume	0.12	0.11	0.12	-0.11	0.37	0.49	0.54	1.00	
(9) Whales Netflow	0.14	0.14	0.09	-0.09	0.04	0.00	0.01	0.07	1.00

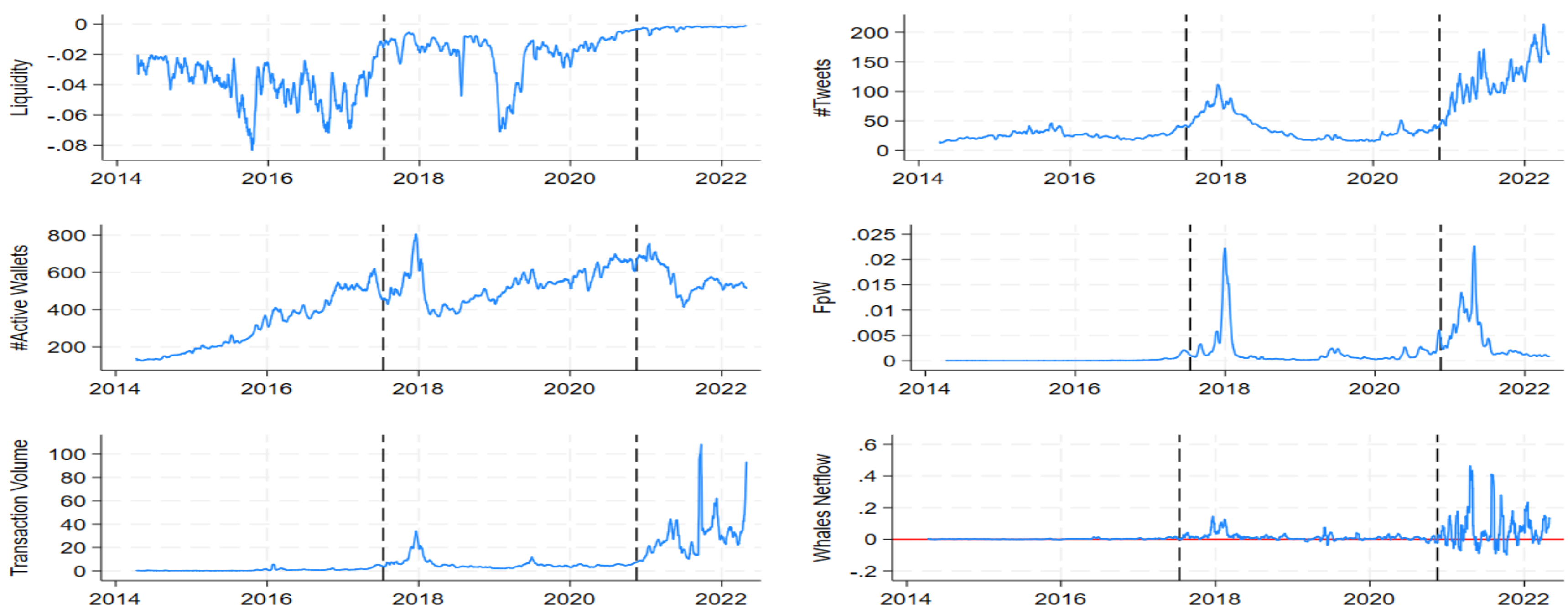

Figure 2: Evolution of number of tweets (#Tweets), number of active wallets (#Active Wallets), transaction fees per weight (FpW), Transaction Volume, and Whales Netflow.

This figure shows the main variables: #Tweets (in thousands), #Active Wallets (in thousands), FpW (in $), Transaction Volume (in $billion), and Whales Netflow (in $billion) from April 10, 2014 to April 30, 2022. The variables are smoothed using a 14-day moving average.

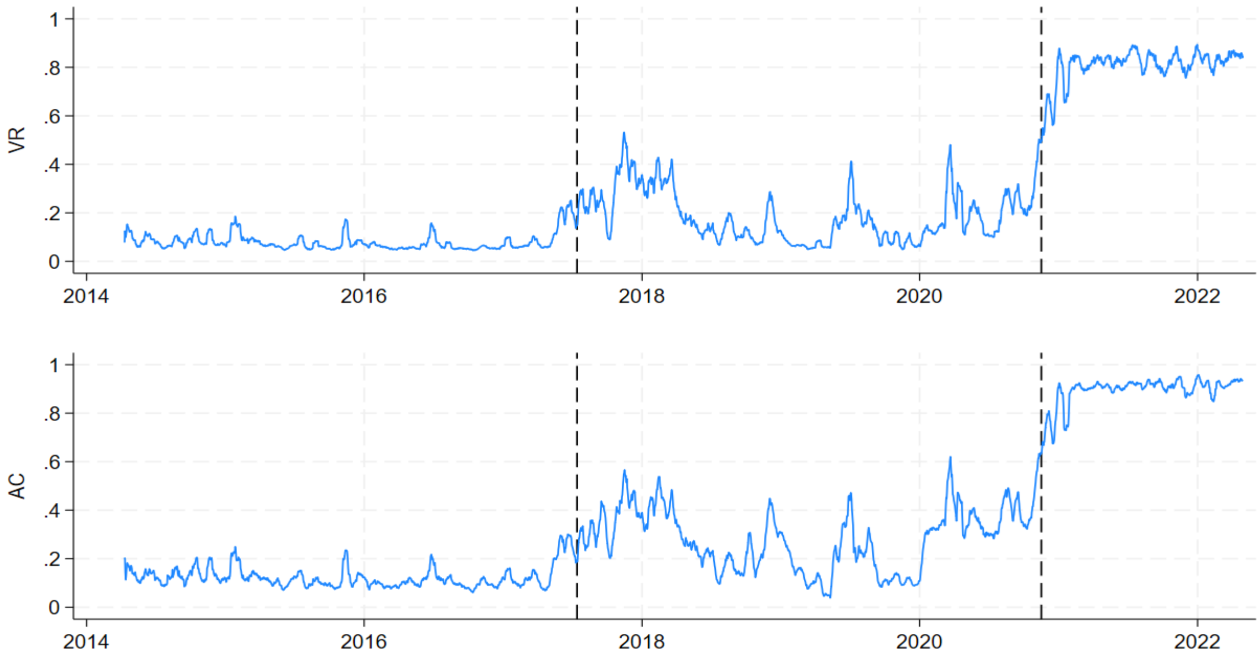

Figure 3: Evolution of Bitcoin Market Efficiency
Daily Bitcoin market efficiency from April 10, 2014 to April 30, 2022. VR and AC are market efficiency measures based on variance ratio test and autocorrelation respectively.

liquidity by 95% in the last period relative to the first period. The number of wallets reached an average of over 565 thousand in period 3, which is 78% more than in period 1. The largest upswings, however, are evident in the transaction fees per weight, transaction volume, and whales' netflow, with an over 20-fold increase in the first and and over 30-fold increase in the latter two.

We also run a by-period correlation for all the variables, however, for brevity provide it in the Appendix III. The results show high correlations amongst our on-chain factors, especially in the periods 2 and 3. Therefore, for consistency we proceed to estimate our models using the on-chain factors separately.

To further test whether the factors behave differently across the three time periods, we first add dummy variables to our baseline model, indicating respectively the time periods separated by the two date breaks. Specifically dummy Period2 acquires a value of 1 for the period between July 14, 2017 and November 16, 2020, and zero otherwise, and dummy Period3 acquirers a value of 1 for the period starting November 16, 2020 onwards, and zero otherwise. To further isolate the potential different effects of our explanatory variables, in the second step we report our analysis for each period separately.

Table 3: Univariate test results.
This table reports the univariate tests for all variables, based on the raw value (non de-trended), used in the empirical analysis before detrending across the three identified distinct periods. VR and AC are market efficiency measures based on variance ratio test and autocorrelation respectively. Number of wallets (#Active Wallets) is in thousand, transaction fees per weight (FpW*) is in USD, and Transaction Volume* and Whales Netflow are in billions of USD. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020, and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	Period 1		Period 2		Period 3		Period 2 - Period 1 (t-value)	Period 3 - Period 2 (t-value)	Period 3 - Period 1 (t-value)
	Mean	SD	Mean	SD	Mean	SD			
VR	0.08	0.06	0.19	0.004	0.81	0.11	0.11*** (-22.88)	0.62*** (-84.79)	0.73*** (-170.00)
AC	0.12	0.08	0.29	0.16	0.90	0.09	0.16*** (-30.85)	0.61*** (-81.09)	0.77*** (-190.00)
Liquidity	-0.04	0.02	-0.02	0.02	-0.002	0.002	0.02*** (-29.84)	0.02*** (-24.04)	0.04*** (-44.48)
#Tweets*	26,127.95	8,096.42	37,409.99	22,273.30	122,245.8	48,080.62	11,282.04*** (-16.43)	84,835.81*** (-50.46)	96,117.85*** (-66.92)
#Active Wallets*	316.81	144.56	534.47	106.17	565.47	96.47	217.66*** (-42.22)	31.00*** (-5.77)	248.66*** (-36.2)
FpW*	0.0002	0.0004	0.002	0.003	0.004	0.005	0.001*** (-15.49)	0.002*** (-12.30)	0.004*** (-27.91)
Transaction Volume*	0.94	1.40	5.78	5.05	30.78	31.42	4.84*** (-31.91)	25.00*** (-27.03)	29.84*** (-32.72)
Whales Netflow	0.002	0.007	0.01	0.06	0.06	0.36	0.01*** (-7.06)	0.05*** (-4.64)	0.06*** (-5.90)
Observations	1,191		1,221		531				

3. Results and Discussion

3.1. OLS Regression Results

Tables 4 through 6 present the regression results with number of wallets, transaction fees per weight and transaction volume as the main on-chain explanatory variable, respectively. The dependent variable in each regression is Bitcoin market efficiency based on variance ratio test (VR)[8]. Column (1) reports the baseline regression results for the full sample period, columns (2), (3), and (4) present the results for each period separately.

All three tables show that liquidity has a significant and positive effect on Bitcoin market efficiency in all periods. This is consistent with the previous findings and supports the theory that increased liquidity facilitates arbitrage activities, and subsequently makes price deviation converge faster towards the efficient price (Chordia et al., 2008; Rösch et al., 2017; Al-Yahyaee et al., 2020). Unreported tests show that the positive liquidity effect significantly increase over time. In fact, the effect in the third period is approximately 20 times greater than that in the first period.

Investor attention has a significant positive effect on Bitcoin market efficiency in periods 1 and 2, providing support for our Hypothesis 1. This is consistent with Vozlyublennaia (2014) findings in various traditional financial markets and support the view that increased attention facilitates information incorporation, thus improving market efficiency. However, in period 3, investor atten-

[8]The correlation between VR and AC is 0.97 and the regression results are quantitatively similar. Therefore, for brevity, we present the results for AC in Appendix IV.

tion appears to lose its positive impact on market efficiency due to the unique dynamics of the Bitcoin bullish period, influenced by the COVID-19 pandemic and heightened market hype at the end of 2020 (Biktimirov and Biktimirova, 2023). During this time, market sentiment, as discussed by Long et al. (2024) and Osman et al. (2024), plays an important role in driving price movements. While increased investor attention typically enhances information flow, it can also introduce noise, as suggested by Da et al. (2011). In this bullish phase, the speculative trading fueled by sentiment likely amplified the noise component, diminishing the overall positive impact of investor attention on market efficiency. Consequently, investor attention's contribution became insignificant.

Finally, from the three tables (columns (1)) we see that Bitcoin market efficiency increases with each subsequent time period. These findings are in line with Urquhart (2016) and Kakinaka and Umeno (2022), who show that Bitcoin market efficiency has increased over time. As an asset matures, this increase in efficiency has also been observed in other asset classes. Smith (2012) find that Greece stock market has a tendency of being more efficient from 2000 to 2009. During this period, Greece was promoted from Advanced Emerging to Developed market by FTSE and S&P. Furthermore, Urquhart (2017) find that from January 1987 to September 2014, while the overall market efficiency of gold and silver does not show significant trends, that of platinum, which has a much shorter history than gold and silver as an investment[9], improves as it develops. This pattern is logical because, with new assets, investors need time to learn and understand them; and as this learning process unfolds, market efficiency improves (Pástor and Pietro, 2003).

Moving onto on-chain factors we first draw focus on the number of active wallets, representing the number of active users. Whilst the full sample analysis indicates a negative effect of active wallets on Bitcoin market efficiency, subsequent sub-period analyses results show that the effect and its sign is period dependent. Specifically, a positive effect is found in period 1, supporting our first hypothesis, but becomes negative in the more mature periods. This shift in the relationship can be explained by the change in Bitcoin user composition. During period 1, which ranges from April 10, 2014, to July 13, 2017, illegal users constitute a significant proportion of Bitcoin users (Foley et al., 2019). Illegal users mainly use Bitcoin as a payment method, therefore, they are intrinsic users because payment system is one of the most important compositions of Bitcoin's intrinsic value (Biais et al., 2023). Therefore, the increase in the number of active users can provide liquidity and facilitate information incorporation, which leads to its association with improved market efficiency. After this period, the proportion of illegal users has dropped greatly as investors' and speculators' interest increases (Foley et al., 2019; Makarov and Schoar, 2021), and Bitcoin market becomes dominated by retail traders, who usually behave irrationally (Kogan et al., 2023). Therefore, the increased number of active users can bring excess noise which impedes information incorporation, which in turn is consistent with the impact of Robinhood investors in the stock market (Eaton et al., 2022).

[9] https://www.lbma.org.uk/publications/the-otc-guide/precious-metal-benchmarks

Table 4: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR). The key independent variables are liquidity, number of tweets (#Tweets), number of active wallets (#Active Wallets), and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR			
	(1)	(2)	(3)	(4)
Liquidity	2.41***	1.22***	3.47***	28.67***
	(0.12)	(0.10)	(0.23)	(4.74)
#Tweets	0.08***	0.05***	0.21***	-0.001
	(0.01)	(0.01)	(0.02)	(0.02)
#Active Wallets	-0.65***	0.33**	-1.53***	-1.13***
	(0.16)	(0.15)	(0.27)	(0.30)
Whales Netflow	0.03*	1.42***	0.18**	0.02*
	(0.01)	(0.39)	(0.09)	(0.01)
Period2	0.06***			
	(0.00)			
Period3	0.63***			
	(0.01)			
Constant	0.18***	0.12***	0.25***	0.88***
	(0.01)	(0.00)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.86	0.24	0.29	0.17

The results for regressions with transaction fees and on-chain transaction volume are reported in Tables 5 and 6, respectively. They show that the effects are positive and significant in period 1, which supports our Hypotheses 2 and 3, whilst insignificant in other periods. Findings in period 1 are consistent with the fact that transaction fees are indicators for Bitcoin blockchain congestion. Specifically, higher rates of transaction activities cause increased congestion, subsequently raising the fees (Easley et al., 2019; Huberman et al., 2021), whereas “No delays imply no fees” (Huberman et al., 2021). Therefore, both increased transaction fees and transaction volume have positive effect on market efficiency. However, blockchain-based settlement has inherent speed limitations, making it unsuitable for high-frequency trading[10]. According to Hautsch et al. (2024), the average waiting time until the transaction is validated from 2018 to 2019 is 41 minutes, and this settlement latency stimulates the development of centralized exchanges. After period 1 (July 2017), exchanges started to expand quickly and this trend continued towards the end of period 3 (2022) (Jahanshahloo et al., 2023). Therefore, the direct effects of transaction fees and on-chain transaction volume on market efficiency becomes insignificant in periods 2 and 3.

Table 5: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR). The key independent variables are liquidity, number of tweets (#Tweets), transaction fees per weight (FpW), and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR			
	(1)	(2)	(3)	(4)
Liquidity	2.43***	1.23***	3.67***	27.40***
	(0.12)	(0.10)	(0.23)	(4.97)
#Tweets	0.06***	0.04***	0.17***	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)
FpW	0.11**	0.17**	0.06	-0.03
	(0.05)	(0.07)	(0.07)	(0.07)
Whales Netflow	0.03*	1.17***	0.19**	0.02
	(0.01)	(0.38)	(0.09)	(0.01)
Period2	0.06***			
	(0.00)			
Period3	0.64***			
	(0.01)			
Constant	0.17***	0.12***	0.25***	0.88***
	(0.01)	(0.00)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.86	0.25	0.27	0.15

For whales netflow, regression results from Tables 4 to 6 show that its increase enhances market efficiency throughout the sample period. A positive netflow indicates that whales are withdrawing

[10]The Bitcoin protocol, by design, requires miners to solve and append a block approximately every 10 minutes, and the block size is limited to 1MB. As a result, transaction processing is slow.

Table 6: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR). The key independent variables are liquidity, number of tweets (#Tweets), transaction volume, and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR			
	(1)	(2)	(3)	(4)
Liquidity	2.39***	1.10***	3.65***	26.95***
	(0.12)	(0.09)	(0.23)	(5.06)
#Tweets	0.06***	0.04***	0.17***	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)
Transaction Volume	0.08*	0.23***	0.07	-0.12
	(0.04)	(0.04)	(0.11)	(0.08)
Whales Netflow	0.03*	1.19***	0.19**	0.02*
	(0.01)	(0.36)	(0.09)	(0.01)
Period2	0.06***			
	(0.00)			
Period3	0.64***			
	(0.01)			
Constant	0.17***	0.12***	0.25***	0.88***
	(0.01)	(0.00)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.86	0.27	0.27	0.16

Bitcoin from exchanges to hold, while a negative netflow suggests they are preparing to sell Bitcoin back to exchanges (Hoang and Baur, 2022). Therefore, the results imply that when more whales, as a proxy for institutional investors, choose to hold Bitcoin, market efficiency improves. This is consistent with Boehmer and Kelley (2009) and Kacperczyk et al. (2021) findings in the stock market.

Overall, we show that only in period 1 do we fully support our direct effect hypotheses, while period 2 shows positive effect for the number of tweets only, and the results for period 3 do not support any of the hypothesized relationships. In period 1, which is from 2014 to 2017, Bitcoin market was still developing, and as a relatively new technology and financial asset, users mainly focused on its transaction function (Foley et al., 2019). Therefore, blockchain-based activities had significant impact on its price efficiency. In the later periods, the market expanded greatly, which drew increased media coverage, attracted broader participation from retail investors and speculators (Smales, 2022; Kogan et al., 2023), which in turn was followed by increased institutional adoption (Huang et al., 2022) and regulatory scrutiny[11]. With the growing accessibility of exchanges (Barbon and Ranaldo, 2024), the impact of blockchain-based activities on price efficiency reduced.

3.2. Mediation Analysis Results

Whilst the previous set of analyses documents the direct effects of on-chain factors, we further argue that these effects may not fully capture the total effect they have on Bitcoin market efficiency. Specifically, through the impact on the off-chain factors (i.e. the mediators), the total effect may in fact be misrepresented. Table 7 reports the direct and indirect impact of on-chain factors on market efficiency[12], as well as their respective contributions to the total effect.[13] [14]

Panel A reports the impact of active users on Bitcoin market efficiency throughout the three periods. In the first period the number of active users has a positive direct impact on market efficiency that accounts for nearly 65%[15] of the total effect. At the same time, increase in the number of active users, is associated with increase in liquidity, which, in line with Eaton et al. (2022) findings, is translated into a positive effect on market efficiency. This indirect effect accounts for more than 30% of the total effect. However, the number of active users does not affect an indirect effect on market efficiency via liquidity during the second and third periods.

Although the indirect effect of the number of active users through the number of Tweets is insignificant in the first period, this relationship changes in the later periods. Specifically, in period 2, the number of active users, in line with Vozlyublennaia (2014) findings, improved market efficiency by increasing investor attention (#Tweets). However, the direct effect of active users, in line with Bloomfield et al. (2009) findings, is associated with a reduction in market efficiency. Therefore, the combined positive effect of active users through investor attention channel and the created noise in the market via the direct channel in this period is negligible and insignificant. In the third period, increase in the number of active users has only a direct negative impact on market efficiency.

Panels B and C report the mediation analysis results for transaction fees per weight and

[11]See Adrian et al. (2022), Iyer (2022), and https://www.coindesk.com/policy/2022/09/20/crypto-needs-global-regulatory-framework-imf-says/

[12]The reported results are based on VR. The results for AC are quantitatively the same and presented in the Appendix IV.

[13]For brevity, we exclude the results for the regressions for the on-chain factor effects on the mediators, and present them in the Appendix V.

[14]The direct and indirect effect are the coefficient of path c and the product of the coefficients of path a and b in Figure 1, respectively.

[15]The % of total is calculated by dividing the direct effect to the sum of the direct and indirect effects. (i.e., 0.33/(0.33+0.15+0.02) * 100).

transaction volume, respectively. For period 1 the direct and indirect effects are positive and significant, with the former playing a more substantial role in the total effect. In period 2 the total effect remains positive and significant, however, the indirect effect is now more pronounced. The final period shows a reversal in the results and our a-priori expectations. Not only does the direct effect become insignificant, but also the indirect effects become negative. In other words, the total effect of transaction fees and volume on Bitcoin market efficiency are negative due to their negative impact on liquidity.

Again, the results highlight the significant differences across the three sub-periods studied. Period 1 supports our theoretical predictions and raised hypotheses that on-chain factors have a positive effect on market efficiency not only directly, but also indirectly, by affecting the off-chain factors, liquidity and investor attention. Period 2 is a more complex one. Apart from the number of active users, the indirect effects are now more pronounced for transaction fees per weight and volume, and the positive overall effect remains.

Finally, in period 3, the direct effects of on-chain factors align with those in period 2, confirming the reduced impact of on-chain factors on market efficiency. However, transaction fees per weight and transaction volume show reversed signs in the total effects on market efficiency, which is driven by their negative relationship with liquidity. This negative relationship between liquidity and transaction activity is also observed in other financial markets. For instance, Yeyati et al. (2008) document that trading volume is associated with decreased liquidity for a set of emerging stock markets during crisis periods. They explain that this occurs when investors rush to liquidate their holdings, triggering a surge in trading activity. This heightened trading demand leads to illiquidity in the market. In Bitcoin market, positive shocks affect the market dynamics more than negative shocks due to "fear of missing out" of retail investors (Baur and Dimpfl, 2018). Period 3 is in Bitcoin bull cycle, prompting investors to flood the market. This surge in trading activity results in a sharp drop in liquidity, resembling patterns observed in the stock market.

4. Conclusion

Since its creation in 2009, Bitcoin has gradually become one of the most studied markets by regulators, academics, and practitioners across the globe. As the need to understand the Bitcoin market grew, so did the development of the links to existing theories and traditional markets. In this paper we contribute to the debate of the evolution of Bitcoin market efficiency and its determining factors. More specifically, we study how a set of off-chain (liquidity and investor attention) and on-chain (number of active users, transaction fees, and transaction volume) factors affect Bitcoin market efficiency throughout time.

First, we show that Bitcoin market efficiency changed significantly throughout our studied period between April 10, 2014 and April 30, 2022. In fact, we show that there are two structural breaks in the time series, characterized by growing levels of market liquidity, investor attention,

Table 7: Mediation analysis results.
Panel A, B, and C report the direct, indirect effects and corresponding percentage of total effects of number of active wallets (#Active Wallets), transaction fees per weight (FpW), and Transaction Volume, respectively, on Bitcoin market efficiency in period 1, 2, and 3. The dependent variable is Bitcoin market efficiency based on variance ratio (VR), independent variables are #Active Wallets, FpW, and Transaction Volume, and the mediators are Liquidity and #Tweets. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022.

Panel A. #Active Wallets						
	Period 1		Period 2		Period 3	
	(1) Coeffi.	(2) % of total	(3) Coeffi.	(4) % of total	(5) Coeffi.	(6) % of total
Direct Effect	0.33**	64.58%	-1.53***	390.00%	-1.13***	100.00%
Indirect Effect						
Liquidity	0.15**	31.25%	-0.08	20.00%	0.01	-0.91%
#Tweet	0.02	6.25%	1.23***	-312.50%	-0.01	0.91%
Total Effect	0.50***		-0.38		-1.12***	
Obs.	1190		1221		531	
Panel B. FpW						
	Period 1		Period 2		Period 3	
	(1) Coeffi.	(2) % of total	(3) Coeffi.	(4) % of total	(5) Coeffi.	(6) % of total
Direct Effect	0.17**	69.70%	0.06	18.18%	-0.03	7.69%
Indirect Effect						
Liquidity	0.03*	18.18%	0.02	6.06%	-0.21***	80.77%
#Tweet	0.03***	12.12%	0.24***	75.76%	-0.03	11.54%
Total Effect	0.23***		0.32***		-0.27***	
Obs.	1190		1221		531	
Panel C. Transaction Volume						
	Period 1		Period 2		Period 3	
	(1) Coeffi.	(2) % of total	(3) Coeffi.	(4) % of total	(5) Coeffi.	(6) % of total
Direct Effect	0.23***	64.10%	0.07	11.76%	-0.12	36.67%
Indirect Effect						
Liquidity	0.10***	28.21%	0.12***	17.65%	-0.18***	60.00%
#Tweet	0.02***	7.69%	0.47***	70.59%	-0.01	3.33%
Total Effect	0.35***		0.66***		-0.31***	
Obs.	1190		1221		531	

market participation, and fees. As such, we proceed with our analysis separately for the three distinct periods. We find that both, on-chain and off-chain factors have a positive effect on Bitcoin market efficiency in period 1, when Bitcoin is in its development stage, which supports our hypotheses and initial theoretical conjectures. In period 2, which can be described as Bitcoin's growth stage, the effect of investor attention remains positive, while the impact of the number of active users becomes negative. In this period, transaction fees and volume loose their significance. Period 3 is the most distinct one, as the effect of investor attention becomes insignificant, refuting all of our hypotheses raised. In addition, we find increased institutional investor participation improves Bitcoin market efficiency across the periods.

In the mediation analysis we acknowledge that the on-chain factor effects on Bitcoin market efficiency could be mediated by the off-chain factors, yielding different total effects. In fact, we show that the overall effects are positive in the first two periods, whilst negative in the last period studied. As Bitcoin market reached an additional development stage, the overall activity (number of users and transaction volume) and associated costs (transaction fees) also increased. This caused a chain of events, characterized by increase in informed traders, off-chain transactions, excessive noise, more regulatory attention, growth and increase of liquidity amongst others. As a consequence, the complex interactions amongst these factors, and the reduced reliance on the on-chain transactions, changed the nature and significance of the on-chain factors on Bitcoin market efficiency.

This study contributes to the broader understanding of market efficiency by extending traditional frameworks to the context of decentralized and rapidly evolving cryptocurrency assets. By examining both on-chain and off-chain factors, the paper bridges the gap between blockchain-specific dynamics and classical financial market theories, offering a framework to study the market dynamics of other blockchain-based assets. The different patterns across the three development stages provide guidance for analyzing other emerging markets and technically disruptive tools in finance. While our sample focuses on a brief period of increased institutional investor adoption, further research could explore its long-term impact and potential to alter the observed relationships.

References

Aalborg, H. A., P. Molnár, and J. E. de Vries (2019). What can explain the price, volatility and trading volume of bitcoin? *Finance Research Letters 29*, 255–265.

Adrian, T., T. Iyer, and M. S. Qureshi (2022). Crypto prices move more in sync with stocks, posing new risks. *IMF Blog*.

Al-Yahyaee, K. H., W. Mensi, H.-U. Ko, S.-M. Yoon, and S. H. Kang (2020). Why cryptocurrency markets are inefficient: The impact of liquidity and volatility. *The North American Journal of Economics and Finance 52*, 101168.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets 5*(1), 31–56.

Augustin, P., A. Rubtsov, and D. Shin (2023). The impact of derivatives on spot markets: Evidence from the introduction of bitcoin futures contracts. *Management Science 69*(11), 6752–6776.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47–78.

Barber, B. M. and T. Odean (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies 21*(2), 785–818.

Barbon, A. and A. Ranaldo (2024, March 29). On the quality of cryptocurrency markets: Centralized versus decentralized exchanges. University of St.Gallen, School of Finance Research Paper Forthcoming, Swiss Finance Institute Research Paper No. 22-38, Available at SSRN: https://ssrn.com/abstract=3984897.

Baur, D. G. and T. Dimpfl (2018). Asymmetric volatility in cryptocurrencies. *Economics Letters 173*, 148–151.

Biais, B., C. Bisiere, M. Bouvard, C. Casamatta, and A. J. Menkveld (2023). Equilibrium bitcoin pricing. *The Journal of Finance 78*(2), 967–1014.

Bianchi, D. and M. Babiak (2022). On the performance of cryptocurrency funds. *Journal of Banking & Finance 138*, 106467.

Biktimirov, E. N. and L. E. Biktimirova (2023). All topics are not created equal: Sentiment and hype of business media topics and the bitcoin market. *Economics Letters 231*, 111311.

Bloomfield, R., M. O'hara, and G. Saar (2009). How noise trading affects markets: An experimental analysis. *The Review of Financial Studies 22*(6), 2275–2302.

Boehmer, E. and E. K. Kelley (2009). Institutional investors and the informational efficiency of prices. *The Review of Financial Studies 22*(9), 3563–3594.

Brauneis, A. and R. Mestel (2018). Price discovery of cryptocurrencies: Bitcoin and beyondwei2018liquidity. *Economics Letters 165*, 58–61.

Brauneis, A., R. Mestel, R. Riordan, and E. Theissen (2021). How to measure the liquidity of cryptocurrency markets? *Journal of Banking & Finance 124*, 106041.

Brauneis, A., R. Mestel, R. Riordan, and E. Theissen (2022). Bitcoin unchained: Determinants of cryptocurrency exchange liquidity. *Journal of Empirical Finance 69*, 106–122.

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

Chen, H. and B.-H. Hwang (2022). Listening in on investors' thoughts and conversations. *Journal of Financial Economics 145*(2), 426–444.

Chordia, T., R. Roll, and A. Subrahmanyam (2008). Liquidity and market efficiency. *Journal of financial Economics 87*(2), 249–268.

Cipriani, M., A. Guarino, and A. Uthemann (2022). Financial transaction taxes and the informational efficiency of financial markets: a structural estimation. *Journal of Financial Economics 146*(3), 1044–1072.

Colliard, J.-E. and P. Hoffmann (2017). Financial transaction taxes, market composition, and liquidity. *The Journal of Finance 72*(6), 2685–2716.

Comerton-Forde, C. and T. J. Putniņš (2015). Dark trading and price discovery. *Journal of Financial Economics 118*(1), 70–92.

Corwin, S. A. and P. Schultz (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *The journal of finance 67*(2), 719–760.

Cramér, H. (1999). *Mathematical methods of statistics*, Volume 26. Princeton university press.

Da, Z., J. Engelberg, and P. Gao (2011). In search of attention. *The journal of finance 66*(5), 1461–1499.

Easley, D., M. O'Hara, and S. Basu (2019). From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics 134*(1), 91–109.

Eaton, G. W., T. C. Green, B. S. Roseman, and Y. Wu (2022). Retail trader sophistication and stock market quality: Evidence from brokerage outages. *Journal of Financial Economics 146*(2), 502–528.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance 25*(2), 383–417.

Foley, S., J. R. Karlsen, and T. J. Putniņš (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *The Review of Financial Studies 32*(5), 1798–1853.

Griffin, J. M., F. Nardari, and R. M. Stulz (2007). Do investors trade more when stocks have performed well? evidence from 46 countries. *The Review of Financial Studies 20*(3), 905–951.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *The American economic review 70*(3), 393–408.

Hautsch, N., C. Scheuch, and S. Voigt (2024). Building trust takes time: limits to arbitrage for blockchain-based assets. *Review of Finance*, rfae004.

Hoang, L. T. and D. G. Baur (2022). Loaded for bear: Bitcoin private wallets, exchange reserves and prices. *Journal of Banking & Finance 144*, 106622.

Huang, X., J. Lin, and P. Wang (2022). Are institutional investors marching into the crypto market? *Economics Letters 220*, 110856.

Huberman, G., J. D. Leshno, and C. Moallemi (2021). Monopoly without a monopolist: An economic analysis of the bitcoin payment system. *The Review of Economic Studies 88*(6), 3011–3040.

Iacobucci, D., N. Saldanha, and X. Deng (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of consumer psychology 17*(2), 139–153.

Iyer, T. (2022). *Cryptic connections: spillovers between crypto and equity markets*. International Monetary Fund.

Jahanshahloo, H., F. Irresberger, and A. Urquhart (2023). Bitcoin under the microscope. *The British Accounting Review*, 101237.

Kacperczyk, M., S. Sundaresan, and T. Wang (2021). Do foreign institutional investors improve price efficiency? *The Review of Financial Studies 34*(3), 1317–1367.

Kakinaka, S. and K. Umeno (2022). Cryptocurrency market efficiency in short-and long-term horizons during covid-19: An asymmetric multifractal analysis approach. *Finance Research Letters 46*, 102319.

Khuntia, S. and J. Pattanayak (2018). Adaptive market hypothesis and evolving predictability of bitcoin. *Economics Letters 167*, 26–28.

Kogan, S., I. Makarov, M. Niessner, and A. Schoar (2023). Are cryptos different? evidence from retail trading. Technical report, National Bureau of Economic Research.

Liu, Y. and A. Tsyvinski (2021). Risks and returns of cryptocurrency. *The Review of Financial Studies 34*(6), 2689–2727.

Lo, A. W. and A. C. MacKinlay (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies 1*(1), 41–66.

Long, S. C., Y. Xie, Z. Zhou, B. M. Lucey, and A. Urquhart (2024). From whales to waves: The role of social media sentiment in shaping cryptocurrency markets. *Ying and Zhou, Zhengyuan and Lucey, Brian M. and Urquhart, Andrew, From Whales to Waves: The Role of Social Media Sentiment in Shaping Cryptocurrency Markets (January 25, 2024)*.

Makarov, I. and A. Schoar (2021). Blockchain analysis of the bitcoin market. Technical report, National Bureau of Economic Research.

Momtaz, P. P. (2024). Decentralized finance (defi) markets for startups: search frictions, intermediation, and the efficiency of the ico market. *Small Business Economics*, 1–33.

Osman, M. B., C. Urom, K. Guesmi, and R. Benkraiem (2024). Economic sentiment and the cryptocurrency market in the post-covid-19 era. *International Review of Financial Analysis 91*, 102962.

Pagnotta, E. S. (2022). Decentralizing money: Bitcoin prices and blockchain security. *The Review of Financial Studies 35*(2), 866–907.

Pástor, L. and V. Pietro (2003). Stock valuation and learning about profitability. *The Journal of Finance 58*(5), 1749–1789.

Rösch, D. M., A. Subrahmanyam, and M. A. Van Dijk (2017). The dynamics of market efficiency. *The review of financial studies 30*(4), 1151–1187.

Scharnowski, S. and H. Jahanshahloo (2025). The Economics of Liquid Staking Derivatives: Basis Determinants and Price Discovery. *Journal of Futures Markets*.

Shen, D., A. Urquhart, and P. Wang (2019). Does twitter predict bitcoin? *Economics Letters 174*, 118–122.

Smales, L. A. (2022). Investor attention in cryptocurrency markets. *International Review of Financial Analysis 79*, 101972.

Smith, G. (2012). The changing and relative efficiency of european emerging stock markets. *The European Journal of Finance 18*(8), 689–708.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology 13*, 290–312.

Sockin, M. and W. Xiong (2023). A model of cryptocurrencies. *Management Science 69*(11), 6684–6707.

Urquhart, A. (2016). The inefficiency of bitcoin. *Economics Letters 148*, 80–82.

Urquhart, A. (2017). How predictable are precious metal returns? *The european journal of finance 23*(14), 1390–1413.

Urquhart, A. (2018). What causes the attention of bitcoin? *Economics Letters 166*, 40–44.

Vlastakis, N. and R. N. Markellos (2012). Information demand and stock market volatility. *Journal of Banking & Finance 36*(6), 1808–1821.

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance 41*, 17–35.

Wei, W. C. (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters 168*, 21–24.

Yeyati, E. L., S. L. Schmukler, and N. Van Horen (2008). Emerging market liquidity and crises. *Journal of the European Economic Association 6*(2-3), 668–682.

Appendix I - Whales Netflow Robustness

The netflow for whales (Whales Netflow*) in Table 8, 9, and 10 is calculated by identifying wallets holding over 1,000 BTC at any point during the day. For these wallets, the netflow is the difference between the Bitcoin received from and sent to identified exchanges in billions of USD.

Table 8: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR) and autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), number of active wallets (#Active Wallets), and the netflow from exchanges to whales (Whales Netflow*). Period2 and Period3 are dummies set to one in periods 2 and 3, respectively, and zero otherwise. Columns (1) and (5) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2) to (4) and (6) to (8) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Liquidity	2.41***	1.18***	3.39***	29.05***	2.72***	1.32***	3.86***	27.18***
	(0.12)	(0.10)	(0.23)	(4.75)	(0.14)	(0.12)	(0.29)	(5.96)
#Tweets	0.08***	0.04***	0.21***	-0.001	0.10***	0.06***	0.23***	0.004
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
#Active Wallets	-0.65***	0.33**	-1.51***	-1.15***	-0.69***	0.46***	-1.85***	-0.60***
	(0.16)	(0.15)	(0.27)	(0.30)	(0.16)	(0.18)	(0.26)	(0.21)
Whales Netflow*	0.03*	1.34***	0.24***	0.02*	0.02*	1.72***	0.26***	0.01
	(0.01)	(0.32)	(0.08)	(0.01)	(0.01)	(0.34)	(0.07)	(0.01)
Period2	0.06***				0.11***			
	(0.00)				(0.01)			
Period3	0.63***				0.66***			
	(0.01)				(0.01)			
Constant	0.18***	0.12***	0.25***	0.88***	0.23***	0.16***	0.35***	0.96***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.26	0.30	0.17	0.87	0.25	0.32	0.25

Table 9: Effects of off-chain factors and transaction fees per weight on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR) and autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), transaction fees per weight (FpW), and the netflow from exchanges to whales (Whales Netflow*). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) and (5) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, columns (2) to (4) and (6) to (8) present the results for each period separately. period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Liquidity	2.43***	1.19***	3.57***	27.69***	2.75***	1.34***	4.07***	26.54***
	(0.12)	(0.10)	(0.23)	(4.99)	(0.14)	(0.12)	(0.30)	(6.08)
#Tweets	0.06***	0.04***	0.16***	-0.01	0.08***	0.06***	0.18***	-0.003
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
FpW	0.11**	0.11	0.05	-0.03	0.09**	0.12	0.02	0.002
	(0.05)	(0.08)	(0.07)	(0.07)	(0.05)	(0.09)	(0.07)	(0.04)
Whales Netflow*	0.02*	1.19***	0.25***	0.02*	0.02*	1.56***	0.28***	0.01
	(0.01)	(0.35)	(0.08)	(0.01)	(0.01)	(0.37)	(0.08)	(0.01)
Period2	0.06***				0.11***			
	(0.00)				(0.01)			
Period3	0.64***				0.66***			
	(0.01)				(0.01)			
Constant	0.17***	0.12***	0.25***	0.88***	0.23***	0.16***	0.35***	0.96***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.26	0.28	0.16	0.86	0.24	0.30	0.24

Table 10: Effects of off-chain factors and transaction volume on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on variance ratio (VR) and autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), transaction volume, and the netflow from exchanges to whales (Whales Netflow*). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) and (5) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, columns (2) to (4) and (6) to (8) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Liquidity	2.39***	1.08***	3.56***	27.25***	2.71***	1.21***	4.06***	26.72***
	(0.12)	(0.09)	(0.23)	(5.09)	(0.14)	(0.11)	(0.30)	(6.09)
#Tweets	0.06***	0.04***	0.16***	-0.01	0.08***	0.06***	0.20***	-0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Transaction Volume	0.08*	0.21***	0.06	-0.13	0.07*	0.25***	-0.17*	0.04
	(0.04)	(0.04)	(0.10)	(0.08)	(0.04)	(0.05)	(0.10)	(0.04)
Whales Netflow*	0.02*	1.11***	0.25***	0.02*	0.02*	1.45***	0.29***	0.01
	(0.01)	(0.30)	(0.08)	(0.01)	(0.01)	(0.31)	(0.08)	(0.01)
Period2	0.06***				0.11***			
	(0.00)				(0.01)			
Period3	0.63***				0.66***			
	(0.01)				(0.01)			
Constant	0.17***	0.11***	0.25***	0.88***	0.23***	0.16***	0.35***	0.96***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.28	0.28	0.16	0.86	0.27	0.30	0.24

Appendix II - Volatility Robustness

Table 11: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency measured with Equation (3). The key independent variables are volatility, number of tweets (#Tweets), number of active wallets (#Active Wallets), and the netflow from exchanges to whales (Whales Netflow). Period 2 and Period 3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, columns (2), (3) and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volatility	-0.12***	-0.06***	-0.19***	-1.37***	-0.14***	-0.06***	-0.24***	-1.39***
	(0.01)	(0.00)	(0.01)	(0.20)	(0.01)	(0.01)	(0.02)	(0.14)
#Tweets	0.08***	0.04***	0.22***	-0.01	0.10***	0.07***	0.23***	0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
#Active Wallets	-0.65***	0.34**	-1.55***	-0.94***	-0.69***	0.49***	-1.85***	-0.42**
	(0.16)	(0.15)	(0.27)	(0.28)	(0.16)	(0.18)	(0.26)	(0.18)
Whales Netflow	0.03*	1.40***	0.17**	0.01	0.02*	1.93***	0.19**	0.01
	(0.01)	(0.38)	(0.09)	(0.01)	(0.01)	(0.42)	(0.08)	(0.01)
Period2	0.06***				0.10***			
	(0.00)				(0.01)			
Period3	0.63***				0.65***			
	(0.01)				(0.01)			
Constant	0.18***	0.12***	0.26***	0.87***	0.24***	0.17***	0.37***	0.95***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.24	0.30	0.22	0.87	0.23	0.35	0.35

Table 12: Effects of off-chain factors and transaction fees on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency measured with Equation (3). The key independent variables are volatility, number of tweets (#Tweets), transaction fees per weight (FpW), and the netflow from exchanges to whales (Whales Netflow). Period 2 and Period 3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, columns (2), (3) and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volatility	-0.12***	-0.06***	-0.20***	-1.35***	-0.14***	-0.06***	-0.25***	-1.38***
	(0.01)	(0.00)	(0.01)	(0.19)	(0.01)	(0.01)	(0.02)	(0.14)
#Tweets	0.06***	0.04***	0.17***	-0.01	0.08***	0.06***	0.19***	0.001
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
FpW	0.10**	0.16**	0.04	-0.03	0.08*	0.18**	-0.001	-0.00
	(0.05)	(0.07)	(0.07)	(0.07)	(0.05)	(0.09)	(0.07)	(0.04)
Whales Netflow	0.03*	1.16***	0.19**	0.01	0.02*	1.67***	0.21**	0.004
	(0.01)	(0.38)	(0.09)	(0.01)	(0.01)	(0.43)	(0.09)	(0.01)
Period2	0.06***				0.10***			
	(0.00)				(0.01)			
Period3	0.64***				0.66***			
	(0.01)				(0.01)			
Constant	0.17***	0.12***	0.26***	0.87***	0.23***	0.17***	0.37***	0.95***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.25	0.27	0.20	0.87	0.23	0.32	0.35

Table 13: Effects of off-chain factors and transaction volume on Bitcoin market efficiency. This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency measured with Equation (3). The key independent variables are volatility, number of tweets (#Tweets), transaction volume, and the netflow from exchanges to whales (Whales Netflow). Period 2 and Period 3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Columns (1) report the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, columns (2), (3) and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	VR				AC			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Volatility	-0.12***	-0.05***	-0.20***	-1.33***	-0.14***	-0.06***	-0.25***	-1.39***
	(0.01)	(0.00)	(0.01)	(0.18)	(0.01)	(0.01)	(0.02)	(0.14)
#Tweets	0.06***	0.04***	0.17***	-0.01	0.08***	0.06***	0.21***	-0.003
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Transaction Volume	0.08*	0.24***	0.05	-0.12	0.07*	0.29***	-0.18*	0.06
	(0.04)	(0.04)	(0.11)	(0.08)	(0.04)	(0.04)	(0.10)	(0.04)
Whales Netflow	0.03*	1.16***	0.19**	0.01	0.02*	1.65***	0.22**	0.003
	(0.01)	(0.36)	(0.09)	(0.01)	(0.01)	(0.39)	(0.09)	(0.00)
Period2	0.06***				0.10***			
	(0.00)				(0.01)			
Period3	0.63***				0.66***			
	(0.01)				(0.01)			
Constant	0.17***	0.12***	0.26***	0.87***	0.23***	0.16***	0.37***	0.95***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531	2942	1190	1221	531
Adjusted R^2	0.86	0.28	0.27	0.21	0.87	0.25	0.32	0.35

Appendix III - Correlation Matrix by Periods

Table 14: Correlation Matrix
This table reports the correlations between measures used in the empirical analysis for the three periods. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020, and period 3 is from November 16, 2020 to April 30, 2022.

	Panel A. Period 1								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) VR	1.00								
(2) AC	0.84	1.00							
(3) Liquidity	0.44	0.41	1.00						
(4) Volatility	-0.44	-0.40	-0.95	1.00					
(5) #Tweets	0.25	0.27	0.21	-0.21	1.00				
(6) #Active Wallets	0.09	0.10	0.07	-0.06	0.03	1.00			
(7) FpW	0.18	0.18	0.08	-0.10	0.13	0.00	1.00		
(8) Transaction Volatility	0.32	0.30	0.25	-0.23	0.14	0.34	0.39	1.00	
(9) Whales Netflow	0.22	0.23	0.13	-0.13	0.10	-0.02	0.28	0.16	1.00
	Panel B. Period 2								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) VR	1.00								
(2) AC	0.86	1.00							
(3) Liquidity	0.43	0.45	1.00						
(4) Volatility	-0.43	-0.49	-0.94	1.00					
(5) #Tweets	0.37	0.38	0.22	-0.21	1.00				
(6) #Active Wallets	-0.04	-0.06	-0.02	0.02	0.37	1.00			
(7) FpW	0.17	0.15	0.03	-0.05	0.46	0.49	1.00		
(8) Transaction Volatility	0.21	0.17	0.10	-0.11	0.53	0.67	0.70	1.00	
(9) Whales Netflow	0.13	0.13	0.06	-0.07	0.10	-0.01	0.04	0.07	1.00
	Panel C. Period 3								
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) VR	1.00								
(2) AC	0.77	1.00							
(3) Liquidity	0.40	0.49	1.00						
(4) Volatility	-0.45	-0.60	-0.88	1.00					
(5) #Tweets	-0.21	-0.22	-0.44	0.38	1.00				
(6) #Active Wallets	-0.14	-0.09	0.00	0.04	0.20	1.00			
(7) FpW	-0.14	-0.14	-0.29	0.24	0.45	0.29	1.00		
(8) Transaction Volatility	-0.17	-0.10	-0.26	0.23	0.38	0.45	0.57	1.00	
(9) Whales Netflow	0.04	0.02	-0.02	-0.01	0.01	0.06	0.05	0.10	1.00

Appendix IV - OLS Regression and mediation analysis for market efficiency based on autocorrelation (AC).

Table 15: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), number of active wallets (#Active Wallets), and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Column (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	AC			
	(1)	(2)	(3)	(4)
Liquidity	2.72***	1.37***	3.96***	26.98***
	(0.14)	(0.12)	(0.29)	(5.93)
#Tweets	0.10***	0.07***	0.24***	0.004
	(0.01)	(0.01)	(0.02)	(0.01)
#Active Wallets	-0.69***	0.47***	-1.87***	-0.59***
	(0.16)	(0.18)	(0.26)	(0.21)
Whales Netflow	0.02*	1.95***	0.20**	0.01
	(0.01)	(0.43)	(0.08)	(0.01)
Period2	0.11***			
	(0.01)			
Period3	0.66***			
	(0.01)			
Constant	0.23***	0.17***	0.36***	0.96***
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.87	0.23	0.32	0.24

Table 16: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), transaction fees per weight (#FpW), and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Column (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	AC			
	(1)	(2)	(3)	(4)
Liquidity	2.75***	1.39***	4.18***	26.38***
	(0.14)	(0.12)	(0.30)	(6.06)
#Tweets	0.08***	0.06***	0.19***	-0.003
	(0.01)	(0.01)	(0.02)	(0.01)
FpW	0.09**	0.19**	0.02	0.002
	(0.05)	(0.08)	(0.07)	(0.04)
Whales Netflow	0.02*	1.66***	0.22**	0.01
	(0.01)	(0.43)	(0.09)	(0.01)
Period2	0.11***			
	(0.01)			
Period3	0.67***			
	(0.01)			
Constant	0.23***	0.17***	0.36***	0.96***
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.86	0.23	0.29	0.24

Table 17: Effects of off-chain factors and number of active wallets on Bitcoin market efficiency.
This table reports the OLS regression results using daily data, in which the dependent variable is Bitcoin market efficiency based on autocorrelation (AC). The key independent variables are liquidity, number of tweets (#Tweets), transaction volume, and the netflow from exchanges to whales (Whales Netflow). Period2 and Period3 are dummies set to one in period 2 and 3, respectively, and zero otherwise. Column (1) reports the baseline regression results for the full sample period from April 10, 2014 to April 30, 2022, and columns (2), (3), and (4) present the results for each period separately. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	AC			
	(1)	(2)	(3)	(4)
Liquidity	2.71***	1.24***	4.16***	26.59***
	(0.14)	(0.11)	(0.29)	(6.07)
#Tweets	0.08***	0.06***	0.21***	-0.01
	(0.01)	(0.01)	(0.02)	(0.01)
Transaction Volume	0.07*	0.28***	-0.16	0.05
	(0.04)	(0.05)	(0.10)	(0.04)
Whales Netflow	0.02*	1.68***	0.23**	0.01
	(0.01)	(0.40)	(0.09)	(0.01)
Period2	0.11***			
	(0.01)			
Period3	0.66***			
	(0.01)			
Constant	0.23***	0.16***	0.36***	0.96***
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	2942	1190	1221	531
Adjusted R^2	0.86	0.26	0.29	0.24

Table 18: Mediation analysis results.
Panel A, B, and C report the direct, indirect effects and corresponding percentage of total effects of number of active wallets (#Active Wallets), transaction fees per weight (FpW), and Transaction Volume, respectively, on Bitcoin market efficiency in period 1, 2, and 3. The dependent variable is Bitcoin market efficiency based on autocorrelation (AC), independent variables are #Active Wallets, FpW, and Transaction Volume, and the mediators are Liquidity and #Tweets. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022.

Panel A. #Active Wallets						
	Period 1		Period 2		Period 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coeffi.	% of total	Coeffi.	% of total	Coeffi.	% of total
Direct Effect	0.47**	70.15%	-1.87***	316.95%	-0.59***	105.36%
Indirect Effect						
Liquidity	0.17**	25.37%	-0.09	15.25%	0.01	-1.79%
#Tweet	0.04	5.97%	1.37***	-232.20%	0.02	-3.57%
Total Effect	0.67***		-0.59**		-0.56***	
Obs.	1190		1221		531	
Panel B. FpW						
	Period 1		Period 2		Period 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coeffi.	% of total	Coeffi.	% of total	Coeffi.	% of total
Direct Effect	0.19***	70.37%	0.02	6.25%	0.002	-0.95%
Indirect Effect						
Liquidity	0.04*	14.81%	0.02	6.25%	-0.20***	95.24%
#Tweet	0.04***	14.81%	0.28***	87.50%	-0.01	4.76%
Total Effect	0.27***		0.32***		-0.21***	
Obs.	1190		1221		531	
Panel C. Transaction Volume						
	Period 1		Period 2		Period 3	
	(1)	(2)	(3)	(4)	(5)	(6)
	Coeffi.	% of total	Coeffi.	% of total	Coeffi.	% of total
Direct Effect	0.28***	66.67%	-0.16	-29.09%	0.05	-35.71%
Indirect Effect						
Liquidity	0.11***	26.19%	0.13***	23.64%	-0.18***	128.57%
#Tweet	0.03***	7.14%	0.58***	105.45%	-0.01	7.14%
Total Effect	0.42***		0.55***		-0.14**	
Obs.	1190		1221		531	

Appendix V - Regression for On-chain Factors on Mediators

Table 19: Regression for On-chain factors effects on the mediators.
This table reports the OLS regression results using daily data, in which the dependent variables liquidity and number of Tweets (#Tweets), and dependent variables are number of active wallets (#Active Wallets), transaction fees per weight (FpW), transaction volume, and the netflow for whales (Whales Netflow). Panel A, B, and C report the results for Period 1, 2, and 3, respectively. Period 1 is from April 10, 2014 to July 13, 2017, period 2 is from July 14, 2017 to November 16, 2020 and period 3 is from November 17, 2020 to April 30, 2022. ***, **, and * indicate statistical significance at 1%, 5%, and 10% levels and t-statistics are reported in parentheses.

	Panel A. Period 1					
	(1) Liquidity	(2) Liquidity	(3) Liquidity	(4) #Tweets	(5) #Tweets	(6) #Tweets
#Active Wallets	0.12**			0.55		
	(0.05)			(0.52)		
FpW		0.03*			0.67***	
		(0.02)			(0.18)	
Transaction Volume			0.09***			0.53***
			(0.01)			(0.15)
Whales Netflow	0.37***	0.33***	0.26***	3.18***	2.20**	2.55**
	(0.08)	(0.08)	(0.06)	(1.13)	(1.03)	(1.06)
Constant	-0.04***	-0.04***	-0.04***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)

	Panel B. Period 2					
	(1) Liquidity	(2) Liquidity	(3) Liquidity	(4) #Tweets	(5) #Tweets	(6) #Tweets
#Active Wallets	-0.02			5.75***		
	(0.02)			(0.39)		
FpW		0.01			1.47***	
		(0.00)			(0.07)	
Transaction Volume			0.03***			2.81***
			(0.01)			(0.13)
Whales Netflow	0.02***	0.02***	0.01***	0.42***	0.32***	0.24**
	(0.00)	(0.01)	(0.01)	(0.12)	(0.11)	(0.10)
Constant	-0.02***	-0.02***	-0.02***	-0.01**	-0.01*	-0.01*
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)

	Panel C. Period 3					
	(1) Liquidity	(2) Liquidity	(3) Liquidity	(4) #Tweets	(5) #Tweets	(6) #Tweets
#Active Wallets	0.0004			4.50***		
	(0.00)			(0.96)		
FpW		-0.01***			2.35***	
		(0.00)			(0.23)	
Transaction Volume			-0.01***			1.87***
			(0.00)			(0.29)
Whales Netflow	-0.0001	-0.00004	0.007	-0.003	-0.01	-0.02
	(0.00)	(0.00)	(0.00)	(0.05)	(0.04)	(0.05)
Constant	-0.00***	-0.00***	-0.00***	0.16***	0.17***	0.11***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)