JOEEP

e-ISSN: 2651-5318
Journal Homepage: http://dergipark.org.tr/joeep

Araştırma Makalesi ● Research Article

Whales And Retail Attention: Dual Channels of Bitcoin Return Dynamics

Balinalar ve Perakende Yatırımcı Dikkati: Bitcoin Getiri Dinamiklerinde İkili Kanallar

Ahmet Furkan Sak [a,*]

[a] Dr. Öğr. Üyesi, Department of Business, Burdur Mehmet Akif Ersoy University, 15030, Burdur, Türkiye
ORCID: 0000-0002-6713-5773

M A K A L E B İ L G İ S İ

Makale Geçmişi:
Başvuru tarihi: 6 Aralık 2026
Düzeltme tarihi: 24 Mayıs 2025
Kabul tarihi: 21 Haziran 2026

Anahtar Kelimeler:
Bitcoin
Balina Işlemleri
Yatırımcı Dikkati

ÖZ

Bu çalışma, balina işlemlerinin ve yatırımcı dikkatinin Bitcoin getirileri üzerindeki etkilerini incelemektedir. 2018 Ocak–2025 Mart dönemine ait 2.647 günlük gözlem kullanılarak, 1, 7, 14 ve 28 gün sonrası ileriye dönük kümülatif anormal getiriler üzerinden bir olay çalışması yöntemi uygulanmıştır. Veriler, ilk 100 saklama dışı cüzdana ait zincir üstü istatistikler ile normalize edilmiş Google Trends verilerinden elde edilmiştir. Balina faaliyetlerinin ekonomik ve istatistiksel olarak anlamlı etkiler ürettiği, olay günlerinde getirilerin sırasıyla %0,79 (7 gün), %1,66 (14 gün) ve %2,44 (28 gün) daha yüksek olduğu tespit edilmiştir. Dikkat şokları anlık bir etki göstermemekte, daha gecikmeli biçimde 14 ve 28 gün ufuklarında sırasıyla %1,76 ve %2,36 oranında artışlara yol açmaktadır. Bulgular, Bitcoin fiyat oluşumunun büyük yatırımcı işlemleri ve perakende yatırımcı dikkati olmak üzere iki kanaldan gerçekleştiğini göstermektedir.

A R T I C L E I N F O

Article history:
Received: Dec 6, 2025
Received in revised form: May 24, 2025
Accepted: June 21, 2026

Keywords:
Bitcoin
Whale transactions
Investor attention

A B S T R A C T

This study examines how whale transactions and investor attention affect Bitcoin returns. An event-study framework with forward cumulative abnormal returns over 1, 7, 14, and 28 days is applied using 2,647 daily observations from January 2018 to March 2025. Data are obtained from on-chain statistics for the top 100 non-custodial wallets and normalized Google Trends. Whale activity generates statistically and economically significant effects, with event-day returns exceeding non-event days by 0.79% (7 days), 1.66% (14 days), and 2.44% (28 days). Attention shocks adjust more slowly, showing no immediate impact but return increases of 1.76% after 14 days and 2.36% after 28 days. Overall, Bitcoin price formation operates through two channels: large-holder order flow and retail-driven attention.

1. Introduction

Bitcoin is a decentralized cryptocurrency that is stored on a public, append-only blockchain and has no central authority. It was introduced in the 2008 white paper "Bitcoin: A Peer-to-Peer Electronic Cash System" by Satoshi Nakamoto (2008). The transaction history is published on a public ledger that is verifiable by anyone. The decentralized nature of Bitcoin allows it to be transferred without central banking and very quickly across borders. It has been adopted over the years as both a medium of exchange and a store of value (Baur & Dimpfl, 2021). Unlike traditional financial assets, Bitcoin provides pseudonymous, tamper-resistant, and publicly verifiable records of transactions that permit detailed, on-chain data analysis (Conti et al., 2018). Further, Bitcoin trades continuously, 24 hours a day, year-round without centralized pauses (Brandvold et al., 2015). Such transparency and 24/7 trading make Bitcoin structurally distinct from traditional markets, allowing for specific identification of price-related events.

In this ecosystem, a prominent subset of actors is called "whales", typically entities or individuals who hold very large balances (e.g., ≥1,000 BTC). Whale transactions are

* *Sorumlu yazar/Corresponding author*.
e-posta: *afsak@mehmetakif.edu.tr*

Atıf/Cite as: Sak, A.F. (2026). Whales And Retail Attention: Dual Channels of Bitcoin Return Dynamics. *Journal of Emerging Economies and Policy*, 11(1), 465-474.

observable on-chain and are price-affecting, given their size. Large buy/sell transactions change the supply-demand balance, and their public transactional history often triggers sentiment responses (e.g., social media notifications) (Long et al., 2025). Therefore, whale activity is a direct order-flow channel and an indirect sentiment channel, and is a primary driver of return behavior in the short- to medium-horizons.

Investor attention is also an important contributor to cryptocurrency markets. Previous research shows strong dependence and even bidirectional causality between search queries for “bitcoin” and Bitcoin prices and/or volatility (Garcia et al., 2014; Kristoufek, 2015; Urquhart, 2018). Google search volume has subsequently been used as a measure of user interest (Cheah & Fry, 2015; Kristoufek, 2013, 2015). Kristoufek (2013) finds a strong positive correlation between Bitcoin price and search frequency on Google and Wikipedia, while Cheah and Fry (2015) find speculative bubbles coincide with surges in attention. Mai et al. (2018) and Polasik et al. (2015) also stress that Bitcoin price and volume are largely a function of popularity and collective perceptions rather than simply fundamentals. More recently, Kapar and Olmo (2021) demonstrate that increases in online searches significantly predict next-period returns. Collectively, these studies justify referring to sudden bursts in search activity as “attention shocks”, which may precede systematic differences in returns.

Two empirical channels, each based on fundamentally different theoretical roots, have been identified. Whale activity has its theoretical roots in the market microstructure literature related to informed trading. Kyle's (1985) original work illustrates how an informed strategic trader will break down their orders, so that private information is incorporated into prices at a rate determined by cumulative price impact (as opposed to instantaneous). Foster and Viswanathan (1996) extended this line of thinking to include multiple differentially-informed traders who compete across many periods. This setting is more analogous to the continuous order-book of Bitcoin, where several very large wallets trade at the same time. Grossman and Stiglitz (1980) provided a complementary equilibrium-based argument. Informationally efficient prices cannot exist when there is a significant cost associated with obtaining information. This is because informed agents must receive compensation for their costs of collecting information. Combined together, both of these fundamental theories suggest that non-custodial whales, who are identifiable from their on-chain order flow but whose information motive is unknown, should demonstrate an influence on prices that is gradual, persistent, and economically rewarded. Feng et al. (2018) empirically demonstrated that large price movements were preceded by informed trading behavior in the Bitcoin market. Also, Makarov and Schoar (2020) found that about 80 percent of Bitcoin return variance is explained by the common component of signed order flow, consistent with the price impact mechanism. The medium-horizon return premium exhibited after whale inflows in Section 5 supports both of these lines of reasoning.

In contrast to the Whale Channel, the Attention Channel is derived from Merton's (1987) incomplete-information equilibrium model. In this model, investors do not purchase securities that they are unaware of and therefore require higher expected returns due to the lack of knowledge about those securities. As the investor base for a particular security increase, the shadow costs of incompleteness decrease. As a result, demand for the security increases and prices rise accordingly, although the increase in price may be delayed while recognition spreads throughout the population. A study by Bodnaruk and Östberg (2009) provides direct empirical evidence supporting this hypothesis. Da, Engelberg, and Gao (2011) found that attention shocks measured using Google SVI lead to significant positive stock returns over a two-week period with subsequent reversal. Therefore, surges in searches for the term "bitcoin" provide a natural empirical proxy for the degree of investor recognition toward the asset: the search surge represents an increase in retail awareness toward the asset; according to Merton's framework, this should result in increased demand pressure and subsequently positive drift as new investors begin to build position(s) in the asset. The 14- and 28-day horizons over which the attention effect is observed in Section 5 are consistent with this friction-laden adjustment process.

Finally, the Whale Channel and the Attention Channel are not independent. Whale transactions are immediately broadcast via "Whale Alert" feeds and social media accounts that serve as attention amplifiers (Saggu, 2022; Long et al., 2025). Therefore, whale transactions can generate retail recognition. Conversely, increased retail attention toward an asset can amplify the price impact of large trades by inducing additional retail order flow in the same direction; a mechanism that the empirical results in Section 5 were designed to capture. The empirical analysis below uses separate event study analyses to identify each channel separately; however, we revisit the combined mechanism implied by theory in our discussion.

In this paper, an event-based analysis of Bitcoin returns is performed, utilizing two observable signals. On-chain movements of the 100 largest Bitcoin wallets are analyzed to identify “whale purchase” event days, while custodial and exchange addresses are excluded to ensure genuine large-holder activity is recorded. Second, “attention shock” event days are identified using spikes in the Google search activity for Bitcoin as a measure of changing investor attention. By examining these two observable phenomena, the paper directly tests how order-flow shocks and investor attention influence short- and medium-horizon return dynamics.

Whale-oriented studies use different approaches. Simulation-based studies (Shen & Shi, 2025) show that concentrated ownership exacerbates price volatility, while event-based studies (Saggu, 2022) show asymmetric short-run responses to price when large transfers are publicized. Sentiment-based studies (Long et al., 2025) emphasize how whales interact with investor sentiment, while machine learning approaches (Herremans & Low, 2025; Azamjon et al., 2023; Kim et al., 2022) incorporate whale trading signals into forecasting models. These studies are still mainly indirect in nature and do not directly test whether whale purchases contribute to systematic returns.

To our knowledge, no existing research has systematically tested the return effects of whale purchases from the largest 100 on-chain Bitcoin wallets within an event-study framework. Furthermore, although other papers have analyzed "attention shocks," the delayed price impact has not been examined within a unified model. This study contributes to the literature by jointly analyzing whale trading and investor attention shocks in an event study design employing a cumulative abnormal return measure over 1-28-day horizons. By integrating verifiable on-chain data with search-based measures of attention, this paper provides the first systematic evidence of how order-flow and attention jointly contribute to Bitcoin price behavior.

As previewed below, the findings show that whale trading exerts a persistent and substantial effect on Bitcoin returns, as cumulative abnormal returns are significantly higher than for the control group on whale event days at the 7-, 14-, and 28-day horizons (reaching about 2.4% at 28 days). By contrast, attention shocks have a muted short-run response but a more pronounced effect at the 14- and 28-day horizons, indicating a more gradual adjustment of prices through investor attention and sentiment. Taken together, the results illustrate dual market effects conveyed through large whale trades and attention shocks acting as a slower-moving indicator of medium-term returns.

2. Literature Review

Investor attention in cryptocurrency markets has received considerable focus in the literature, primarily using Google search data. Kristoufek (2013) examined potential relationships between Bitcoin prices and Google/Wikipedia searches and found asymmetric effects with respect to trend levels, pointing to attention as a nonlinear variable. Kristoufek (2015) expanded this investigation using wavelet coherence and found that attention metrics were associated with speculative dynamics across horizons, while still being subject to fundamental and technical drivers. Garcia et al. (2014) advanced this perspective using digital traces such as search volumes, user growth, and word-of-mouth communication. They noted self-reinforcing feedback loops associated with periods of heightened attention, where surges in attention often preceded reversals and served as early-warning signals. Glaser et al. (2014) further claimed that search-based attention was mostly driven by speculative motives, while Bouoiyour et al. (2015) showed using frequency-domain Granger causality that speculative demand drove price dynamics. Cheah and Fry (2015) also concluded that speculative dynamics increased volatility. In a related vein, Mai et al. (2018) and Polasik et al. (2015) suggested that the measures of attention, proxied by popularity and perceived value, played a major role in price formation.

Focusing on the Turkish context, Yıldırım (2020) used an ARDL bounds testing framework and found long-run cointegration between Google search interest in "Bitcoin" and Bitcoin/USD prices, while Granger causality testing identified a unidirectional association from price changes to search intensity. Dulupçu, Yiyit, and Genç (2017) found evidence that Google search queries impact Bitcoin's price dynamics in Türkiye and, importantly, indicated that attention-driven feedback effects are not just a feature of global markets. More recently, Uzun (2024) explored Bitcoin trading volumes in Türkiye, and together with Google search data, reported cointegration and causality patterns that are consistent with the literature. These studies thus confirm that attention-based proxies provide valuable insights into Bitcoin dynamics in both developed and emerging markets.

Based on both the global and Turkish body of evidence, later research employed more advanced econometric and computational techniques. Dastgir et al. (2019) applied a copula-based Granger causality in distribution test and uncovered bidirectional connections between searches and returns, particularly in the distribution's tails. Nasir et al. (2019) combined VAR, copula, and non-parametric methods and suggested that shocks to search values drove both returns and trading volume, with effects lasting approximately a week. Bleher and Dimpfl (2019) created multi-year Google search indices and reported limited return predictability but strong volatility predictability, concluding that Google search data may better identify shifts in perceived risk than contemporaneous directional pressure. Mittal et al. (2019) included search intensity and Twitter signals into frameworks based on machine learning (RNN/LSTM), and found that although queries and tweet volumes were strong predictors of prices, textual sentiment added little incremental value. Kapar and Olmo (2021) included Google search variables into a vector error correction model together with fundamentals (S&P 500, gold, financial stress), providing evidence that the efficient price for Bitcoin during 2018-2019 could largely be explained by Google search activity. Finally, Cretarola and Figà-Talamanca (2021) provided a theoretical model of bubble formation in which surging search activity coincided systematically with exuberant episodes such as 2012-2013 and 2017. Taken together, these findings provide further evidence that investor attention has predictive capacity for returns, volumes, volatility, and bubble dynamics for Bitcoin. In addition, Google Trends data need to be carefully normalized and anchored across horizons, since observed values are on a 0-100 scale within rolling windows. This motivates the use of relative changes rather than absolute levels when defining attention shocks (Bleher & Dimpfl, 2019).

In parallel with the attention literature, there is an increasing number of studies on large-holder activity, commonly referred to as "whale transactions." Shen and Shi (2025) used an Artificial Bitcoin Market model to recreate stylized market characteristics of volatility clustering and fat tails, showing that increasing whale ownership from 1% to 6% more than doubled daily volatility. This underscores the destabilizing potential of concentrated holdings. Event-based analyses offer supportive evidence. Saggu (2022) studied intraday returns of Bitcoin relating to changes in Tether supply and discovered asymmetric effects: positive returns resulting from minting events, especially when announced via Whale Alert, whereas burning events had a negligible impact. Long et al. (2025) applied sentiment lexicons, combined with time-varying VAR, and demonstrated that whale activities were amplified by current sentiment, enabling large holders to transmit volatility. From a measurement standpoint, Herremans and Low (2025) noted that net BTC flows in and out of exchanges have a strong correlation with daily volatility. Yet they also pointed out that address tagging is incomplete, custodial wallets are sometimes misclassified, and Whale Alert, among others, misses transaction activity, implying that on-chain aggregates should be treated as probabilistic approximations and combined with event identifiers for analytical reliability.

Recently, machine learning and AI have been applied to whale data. Azamjon et al. (2023) combined on-chain flows with Whale Alert tweets in a Q-learning framework, enabling improved jump predictions versus either data source alone. Kim et al. (2022) used a self-attention-based multi-LSTM model, trained on on-chain features, that accurately predicts the Bitcoin price. Herremans and Low (2025) employed Synthesizer-Transformer models combined with CryptoQuant and Whale Alert events to predict extreme volatility spikes, with whale variables ranking among the top predictors of next-day volatility.

To summarize, two complementary strands emerge. Attention-based proxies, of which Google searches are the clearest case, capture speculative demand driven by sentiment, and provide leading indicators of returns, volatility, and bubbles. Whale activity, in contrast, captures direct order-flow shocks that alter supply-demand balances while also influencing expectations when salient. Importantly, these two channels can interact; for example, notable whale trades can activate spikes in attention, while spikes in attention can intensify the price impact of whale trades.

Nevertheless, although it is an emerging area, the literature is not yet complete. Attention studies usually rely on aggregate correlations or predictive models without isolating specific events, while whale studies focus on volatility, simulations, or sentiment rather than directly quantifying return dynamics. To our knowledge, systematic evidence on how on-chain purchases by the top 100 wallets translate into post-event return differentials remains limited. This paper addresses that gap by implementing an event-study framework that isolates whale purchase days and attention shocks, comparing their forward cumulative abnormal returns against non-event benchmarks. By incorporating verifiable blockchain and Google search data, this study provides systematic evidence on how large-holder purchases and investor attention jointly affect Bitcoin returns.

3. Data And Variables

This paper uses three daily series to identify key influences on the return dynamics of Bitcoin: the daily closing price of Bitcoin (BTC_PRICE), the Google Trends index for the term “bitcoin” (GOOGLE), and net purchases by the largest 100 Bitcoin wallets (WHALE). Each variable provides different perspectives on price formation, ranging from market fundamentals to investor attention and large-holder behavior.

Table 1. Descriptive Statistics

Variable	BTC_PRICE	GOOGLE	WHALE
N	2,647	2,647	2,647
Mean	30,306	0.81	776
Std. Dev.	24,808	0.51	4,339
Min	3,237	0.24	0
Max	106,146	5.55	100,001
Skewness	1.01	2.48	13.69
Kurtosis	0.3	9.73	235.59

The dataset spans the period January 1, 2018, to March 31, 2025, resulting in 2,647 daily observations. The daily closing price of Bitcoin averages USD 30,306, with substantial dispersion (SD = 24,808) ranging from USD 3,237 to USD 106,146. The distribution of daily closing prices is moderately right-skewed (1.01), reflecting larger upward movements above the mean, and a kurtosis of 0.30, suggesting that extreme price movements occurred only slightly more frequently than under a normal distribution.

The Google Trends index, acting as a measure of investor attention, has an average value of 0.81, ranging from 0.24 to 5.55. It also exhibits high skewness (2.48) and high kurtosis (9.73), indicating that market attention is relatively subdued with occasional large spikes in attention during periods of heightened market interest. For the empirical analysis, attention shocks are defined using percentage changes in the Google Trends index rather than raw levels, in order to capture relative surges in search intensity.

Whale activity is based on on-chain statistics for the 100 largest Bitcoin wallets, obtained from BitInfoCharts as of March 31, 2025. Custodial and exchange-tagged addresses were excluded to ensure that the flows reflect genuine large-holder activity rather than intermediary transfers. The daily series is expressed in units of Bitcoin (BTC) and exhibits extreme variation: the mean daily net inflow is 776 BTC, with a highly right-skewed distribution (skewness = 13.69) and very high kurtosis (235.59). This reflects infrequent but extremely large transfers, occasionally reaching approximately 100,000 BTC. Because the data are based on a fixed snapshot, this series is internally consistent;

however, replications at later dates may yield differences due to subsequent transfers or address reclassification. This limitation mainly affects precise magnitudes, not the qualitative direction of the results.

Binary event indicators were subsequently constructed to define abnormal changes. For the primary specification, the whale event was defined as a day with net inflow exceeding 50 BTC into the top 100 wallets, a value near the median of positive net-inflow days over the sample, and the attention event as a day with a 15% increase in Google Trends. For robustness, a stricter whale threshold (>100 BTC) and a stricter attention threshold (>20%) were also considered as alternative definitions. The use of additional definitions ensures that the returns are not driven by an arbitrary cutoff and allows a systematic comparison of pre-event and post-event returns with no-event controls over short- and medium-term horizons.

4. Methodology

The empirical strategy is based on an event-study framework that examines how asset returns behave after well-defined market events. In this study, the events of interest are defined from two observable sources: whale purchases derived from on-chain flows of the 100 largest Bitcoin wallets and attention shocks proxied by sharp increases in the Google Trends index. This approach enables systematic comparisons of return behavior on event days versus non-event days, providing a transparent way to test for changes in market behavior around these shocks.

Daily Bitcoin log returns were computed as:

$$R_t = \ln(P_t) - \ln(P_{t-1}) \quad (1)$$

where P_t denotes the daily closing price. Abnormal returns under the constant-mean-return model are defined as $AR_t = R_t - \mu$, where μ is the sample-mean daily log return (μ = 0.068%) over the estimation window. For each event day t, forward cumulative abnormal returns (CAR) over horizons $h \in \{1,7,14,28\}$ were defined by:

$$CAR_{t,h} = \Sigma_{i=1}^{h} AR_{t+i} \quad (2)$$

These windows align the forward cumulative abnormal returns with the event date and capture short-run and medium-run dynamics. Because events can occur in quick succession, the forward windows can "overlap," but all observations are preserved to provide maximum statistical power. Forward window overlap can induce serial correlation in the $CAR_{t,h}$, so inferences are based on heteroskedasticity-consistent White (HC1) standard errors and are checked against Welch's unequal variance tests. Welch results are used as a robustness check and are omitted from tables for brevity.

Events were established using binary indicators. In the base specification, a whale event ($Event_t^W = 1$) is defined whenever the net inflow into the largest 100 wallets exceeds 50 BTC, a value near the median of positive net-inflow days over the sample, and an attention event ($Event_t^G = 1$) is designated as that day when the Google Trends index increased by more than 15% in a single day. Each event family is treated in isolation in separate regressions (i.e., $Event_t$ corresponds to one definition at a time).

A simple anchoring procedure was used to make the daily Google Trends data consistent across the full sample. Google Trends reports values that adjust relevant time periods back to a 0-100 scale, and raw daily pulls are not comparable across long horizons. The first step was to collect the monthly Google Trends value for every month. Then, within that month, the daily values were adjusted to shares of the monthly total to multiply by the monthly value such that the daily values always summed to match the reported monthly level. For instance, if January's monthly index was 30, and the daily values summed to 1800, then a day with a daily value of 60 was a share of $60/1800 = 0.0333$, which scales to $30 \times 0.0333 = 1.0$. After this adjustment, percentage changes at the daily level were calculated to define attention shocks.

For robustness, stricter criteria were also used (whale net purchases >100 BTC; Google search increases >20%). These alternative thresholds serve as reassurance against arbitrary event cutoffs being an artifact of the findings.

Post-event outcomes were compared to non-event benchmarks using a regression-based difference-in-means framework:

$$CAR_{t,h} = \alpha + \beta \cdot Event_t + \varepsilon_t, \quad (3)$$

where $CAR_{t,h}$ is the forward cumulative abnormal returns over horizon h aligned with event day t, and $Event_t$ is the event dummy. The coefficient β captures the mean difference in forward cumulative abnormal returns between event days and non-event days for the selected horizon and event definition. The parameters are estimated by OLS with heteroskedasticity-robust (White/HC1) standard errors, appropriate for financial data which may exhibit time-varying volatility.

Also, Welch's unequal-variance t-tests (two-sided) were applied to directly compare the event and control groups. This is a complementary test that does not rely on equal variances for the samples as a benchmark of sorts. The outputs reported, therefore, include group means for event and control days; their differences; White/HC1 t-statistics, Welch statistics/p-values, and sample sizes (including the proportion of positive outcomes). This dual inference strategy facilitates transparency and reduces the reliance on a single estimation technique.

Overall, the structure is intended to account for both channels of market impact: whale purchases as direct order-flow shocks and attention shocks as sentiment-driven demand shifts. Combined with blockchain verifiable transaction data and attention measures from search, as well as effects across horizons, the design strikes a balance between statistical rigor and economic interpretability.

5. Findings

This section presents the empirical findings from the event-study analysis. Results are reported in three steps: First, the time-series properties of the data are evaluated through unit root and stationarity tests. Second, diagnostic checks are conducted on the return series. Finally, the main event-study estimates for whale activity and Google search shocks are reported.

Table 2. Unit Root and Stationarity Test Results for Bitcoin Price and Returns

Series	Test	Statistic	p-value	Interpretation
Log Price	ADF	-0.485	0.895	Non-stationary
Log Price	PP	-0.554	0.881	Non-stationary
Log Price	KPSS	6.376	0.010	Non-stationary
Log Return	ADF	-18.181	<0.001	Stationary
Log Return	PP	-53.954	<0.001	Stationary
Log Return	KPSS	0.155	0.100	Stationary

Table 2 summarizes unit root and stationarity tests for Bitcoin log prices and returns. The results indicate that log prices are non-stationary: neither ADF (-0.485, p=0.895) nor PP (-0.554, p=0.881) reject the unit root, while the KPSS test rejects stationarity (6.376, p=0.010). By contrast, the daily log returns are clearly stationary across all three tests: both ADF (–18.18, p<0.001) and PP (–53.95, p<0.001) tests reject the unit root, whereas KPSS (0.155, p>0.10) fails to reject stationarity. This result is fully aligned with financial theory and prior empirical evidence that asset prices follow a random walk, while returns are stationary. Accordingly, the event-study design is based on daily log returns (Campbell, Lo & MacKinlay, 1997; Kothari & Warner, 2007).

Table 3. Diagnostic Test Results for Bitcoin Return Series

Test	Statistic	p-value
Durbin–Watson	1.997	-
Ljung–Box Q (lag=10)	19.95	0.030
ARCH LM (lags=10)	60.43	<0.001
White test	9.11	0.010
Jarque–Bera	22,502	<0.001

Table 3 provides the diagnostic checks of the return series. The Durbin–Watson statistic is close to 2 (1.997), indicating no significant first-order autocorrelation. The Ljung–Box Q–test, at lag 10, gives 19.95 (p=0.03), which shows some mild higher–order dependence, but this is not strong enough to affect inference in larger samples. As expected for financial data, the ARCH LM test strongly rejects homoskedasticity (60.43, p<0.001), thus confirming volatility clustering. The White test (9.11, p=0.01) provides additional evidence for heteroskedasticity, while the Jarque–Bera statistic (22,502, p<0.001) reveals extreme non-normality. These features justify the use of heteroskedasticity-robust (White/HC1) standard errors and complementary reliance on Welch's t-tests, which remain standard practice in event-study designs.

Table 4. Main Event-Study Results (Forward Cumulative Abnormal Returns)

Case	Number of Events	Horizon (days)	Mean Event	Mean Control	Mean Difference	t-stat	p-value
WHALE >50 BTC	975	1	−0.02%	0.01%	−0.03%	−0.22	0.83
WHALE >50 BTC	969	7	0.46%	−0.32%	0.79%	2.14	0.03**
WHALE >50 BTC	964	14	1.02%	−0.64%	1.66%	3.16	<0.001***
WHALE >50 BTC	952	28	1.64%	−0.80%	2.44%	3.19	<0.001***
GOOGLE Δ>15%	331	1	−0.02%	0.00%	−0.02%	−0.10	0.92
GOOGLE Δ>15%	330	7	0.73%	−0.14%	0.87%	1.45	0.15
GOOGLE Δ>15%	329	14	1.51%	−0.25%	1.76%	2.17	0.03**
GOOGLE Δ>15%	327	28	2.15%	−0.21%	2.36%	2.16	0.03**

Notes: ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively. Mean Difference = Mean(Event) − Mean(Control). Abnormal returns are computed under the constant-mean-return model defined in Section 3 (μ = 0.068% per day). The decline in the number of events as the horizon lengthens reflects data availability, since forward return windows are truncated near the end of the sample.

Table 4 reports the core event-study results under the baseline thresholds, where Google Δ>15% refers to days when Google search activity for "Bitcoin" increases by more than 15% relative to the previous day, and Whale (>50 BTC) refers to days when the net inflow into the 100 largest non-custodial wallets exceeds 50 BTC. Several important regularities emerge from the analysis.

Effects are strongest and most persistent for whale activity. On days when the largest 100 wallets record net inflows exceeding 50 BTC, subsequent Bitcoin returns are systematically higher than those observed on non-event days. At the one-day horizon, the difference is statistically insignificant (−0.03%). The impact strengthens over time: 0.79% higher after 7 days, 1.66% higher after 14 days, and 2.44% higher after 28 days compared to non-event days. Given Bitcoin's high volatility, these magnitudes are economically meaningful and suggest that large-holder purchases convey informed-trading signals that the market gradually incorporates. The consistency of these effects across a large number of whale events (approximately 950-975) alleviates small-sample concerns and strengthens inference.

By contrast, shocks to the attention of investors, as indicated by daily Google search activity rising by more than 15%, exhibit a delayed adjustment pattern in the data. At 1- and 7-day horizons, there are no statistically meaningful differences between event and non-event days, suggesting that spikes in retail attention do not exert immediate price pressure. However, the effects emerge at longer horizons: 1.76% higher after 14 days and 2.36% higher after 28 days relative to non-event days. This pattern is consistent with gradual sentiment formation, with effects materializing over medium horizons.

Taken together, the results reveal two complementary mechanisms through which different market participants influence Bitcoin dynamics. Whale trades are high-concentration and high-volume movements that redirect order flows, which lead to durable price adjustments, whereas attention shocks move more diffusely, reflecting the gradual mobilization of retail investors.

Appendix Table A1 presents comparable results under stricter event definitions, where WHALE >100 BTC captures days when the net inflow to the top 100 non-custodial wallets exceeds 100 BTC, and GOOGLE Δ>20% captures days when Google activity related to the search term "Bitcoin" rises by more than 20% relative to the preceding day. The whale measure remains statistically significant at 7–28 day horizons even under this stricter threshold. In contrast, the Google search measure is sensitive to the threshold: Δ>20% shocks are marginally significant at the 7–14 day horizons (at the 10% level) but lose significance at 28 days. This pattern suggests that larger bursts of attention induce faster price adjustment, but lack continuity compared to smaller, more frequent attention shocks.

Overall, the evidence shows that both whale and attention shocks forecast Bitcoin returns. Whale inflows are a timely and reliable signal of informed trading, whereas attention shocks reflect more gradual surges of retail investor sentiment leading to price adjustments. The complementary mechanisms highlight that overall cryptocurrency markets are sustained by institutional-scale order flows first and then by decentralized attention. In addition to the statistical robustness, the findings reveal a practical implication: monitoring of the large-wallet flows provides signals for short to medium-term market movements, while a change in search volume serves as an early-warning indicator of sentiment-driven volatility. Together, the observations show that Bitcoin price formation is a two-channel mechanism; both whales and retail investors are critical yet distinct channels of market impact.

Table 5. Risk-adjusted Performance of the Whale-event Strategy Across Horizons

Metric	Buy & hold	H = 1	H = 7	H = 14	H = 28
Alpha per trade	-	−0.03%	0.79%	1.66%	2.44%
Annualised IR	-	−0.17	0.60	0.63	0.45
Annualised return	24.8%	17.6%	49.3%	51.4%	46.1%
Gross Sharpe	0.29	0.18	0.71	0.75	0.65
Net Sharpe (30 bps)	0.29	−1.38	0.46	0.62	0.58

Notes: Sample 2018-01-01 to 2025-03-31, N = 2,647 days; whale events (>50 BTC threshold) N = 975. IR = per-trade alpha / pooled tracking error, annualised by √(365/H). Sharpe uses rf = 5% and √365. Net Sharpe applies a 30-bps round-trip cost at trade close.

Table 6. Sensitivity of Net Sharpe to Transaction Costs (H = 28)

Round-trip cost	Net annualised return	Net Sharpe
0 bps (gross)	46.1%	0.65
20 bps	43.5%	0.60
30 bps	42.2%	0.58
40 bps	40.9%	0.56

Notes: Annualised volatility 63.8%; approximately 135 event days per year over the sample.

The whale-event premium also has economic implications. Table 5 provides risk-adjusted performance statistics for a long-only event-based strategy where it takes a position in Bitcoin on every whale day and holds the position for H days. The annualized information ratio is at its highest (0.63) when H = 14 and remains at 0.45 when H = 28. Additionally, the gross Sharpe ratios of 0.65-0.75, based upon a horizon of 7 days or longer, are greater than the 0.29 Sharpe of a Bitcoin buy-and-hold during the same time period by more than a factor of two. The sensitivity of the 28-day net Sharpe to a round-trip transaction cost of 20-40 basis points, a very reasonable estimate of the fees charged by tier-1 spot exchanges, is found in Table 6. The 28-day net Sharpe Ratio remains within the interval [0.56, 0.60] over this range of costs. In contrast, the one-day strategy results in a net-negative return under realistic frictional costs and is included in Table 5 only for completeness. The fact that the medium-horizon profitability persists through transaction costs, but disappears at the daily frequency, is consistent with a gradual price-impact mechanism in which information contained in large on-chain orders is reflected in price changes over days rather than minutes.

In order to measure the possible impact of look-ahead bias due to the use of the BitInfoCharts snapshot dated March 31, 2025, we divide our sample into two subsamples: an "early" subsample covering the years 2018 through 2021, and a "late" subsample covering the years 2022 through 2025. The results regarding the whale-event premium are presented in Table 7. It appears that the medium-horizon effects persist in both subsamples. Specifically, the 14-day premium is statistically significant at the 1% level in the early period and at the 5% level in the late period (2.50% and 1.27% respectively); similarly, the 28-day premium is statistically significant at the 1% level in the early period (3.96%) and marginally significant in the late period (1.62%, $p < 0.10$). The higher levels of premiums found in the earlier subsample appear to be generally consistent with the structural evolution of the Bitcoin market during the sample, including increased participation by institutions and increased on-chain liquidity subsequent to 2021, which would reduce the price impact of large wallet inflows. This consistency of sign and significance across the snapshot date demonstrates that the medium-horizon return premium represents a genuine market mechanism rather than some mechanical feature related to how the top-100-wallet set was determined.

Table 7. Robustness to Look-ahead Bias: Subsample Analysis of the Whale-event Premium

Subsample	N (event)	N (control)	H = 1	H = 7	H = 14	H = 28
Early (2018-2021)	364	1,097	0.15% (0.55)	1.33% (2.07)**	2.50% (2.70)***	3.96% (2.86)***
Late (2022-2025)	611	574	−0.18% (−1.09)	0.48% (1.09)	1.27% (2.01)**	1.62% (1.82)*

Notes: Mean event-control differences in cumulative abnormal returns, with Welch t-statistics in parentheses. ***, **, * denote significance at the 1%, 5%, and 10% levels. N (event) shown for H = 1; small declines at longer horizons due to forward-window truncation.

6. Conclusion

This research examines how two observable signals, whale purchases by the largest Bitcoin holders and spikes in investor attention proxied by Google searches, affect Bitcoin's return dynamics. An event-study framework for daily abnormal returns at four horizons (1, 7, 14, and 28 days) yields several robust conclusions.

First, whale behavior has a persistent effect. Net inflows to the top 100 non-custodial wallets are followed by significantly higher forward cumulative abnormal returns at horizons beyond one week. Event days yield returns that are 0.79% higher than non-event days after 7 days, 1.66% higher after 14 days, and 2.44% higher after 28 days. All of these magnitudes are economically meaningful given Bitcoin's high volatility and consistent with whales acting as order-flow signals, which the market incorporates slowly. The persistence across nearly 1,000 events and under a stricter 100-BTC criterion adds to the robustness of findings.

Second, investor attention shocks have a slower but complementary effect. Daily spikes in Google searches above 15% do not cause immediate price reactions. Instead, differences emerge over medium horizons, with returns on event days being 1.76% higher than on non-event days after 14 days and 2.36% higher after 28 days. Robustness checks with a 20% threshold show that stronger attention shocks produce marginally significant effects at the 7-14 day horizons, but persistence decreases. These findings suggest that attention shocks capture sentiment accumulation instead of instantaneous trading pressure, highlighting the important role that retail-driven narratives play in cryptocurrency markets.

Together, these results highlight a dual mechanism of market impact. Whale trades operate through a direct channel, shifting supply-demand balances and signaling informed activity, whereas attention shocks operate through an indirect channel, shaping expectations more gradually and amplifying speculative dynamics. The two forces are not entirely independent: visible whale flows may spark retail attention, while rising attention may magnify the impact of large trades. This feedback loop underscores the interplay between institutional-scale order flow and retail sentiment in cryptocurrency markets.

From a methodological perspective, the study contributes in two ways. First, it uses a clean event-study design based on verifiable on-chain transactions of the top 100 wallets, excluding custodial and exchange-tagged addresses. To our knowledge, this is one of the first applications to directly quantify the return impact of whale purchases via an event-study framework. Second, by integrating whale events with attention shocks across multiple horizons, the analysis bridges two strands of literature, namely large-holder behavior and investor attention, within a unified empirical setting. Inference is strengthened through heteroskedasticity-robust and Welch tests, ensuring robustness against variance heterogeneity.

The findings have several implications. For investors and traders, monitoring whale flows and attention spikes can improve portfolio timing and risk management, given that the effects unfold predictably over one- to four-week horizons. For regulators and policymakers, the evidence highlights the outsized role of large holders in shaping market dynamics and the amplifying effect of retail attention. Potential measures, such as greater transparency in large transfers, on-chain monitoring dashboards, or investor education about retail herding, may help mitigate episodes of excessive volatility.

Naturally, the study has limitations. Whale activity was measured using a fixed snapshot of the top 100 wallets as of March 31, 2025. While this ensures internal consistency, replications may differ due to subsequent transfers or reclassification of addresses. Similarly, Google Trends requires normalization, in particular the use of percentage-change thresholds rather than absolute levels. These issues may affect estimated magnitudes, though not the anticipated direction.

Future studies could extend these findings in several ways. The utilization of continuous measurements of whale intensity (e.g., flows divided by market capitalization) would provide further depth in understanding the effects of trade size. Looking into heterogeneity based on market states (bull or bear periods) in estimation specifications, possibly via using Markov-switching or quantile regressions, may provide evidence of asymmetry. Intraday event studies could capture short-lived responses that daily data may miss. Lastly, cross-asset analysis for Ethereum, stablecoins, or other major cryptocurrencies would enhance external validity and generality.

In sum, the study shows that both whale purchases and investor attention shocks exert systematic and predictable influences on Bitcoin returns, albeit through distinct temporal channels. Integrating blockchain transparency with behavioral indicators offers a foundation for future work on how institutional-scale actions and retail sentiment jointly shape cryptocurrency markets.

References

Azamjon, M., Sattarov, O., & Cho, J. (2023). Forecasting Bitcoin volatility through on-chain and Whale-Alert tweet analysis using the Q-learning algorithm. *IEEE Access*, *11*, 108092–108103. https://doi.org/10.1109/ACCESS.2023.3317899

Baur, D. G., & Dimpfl, T. (2021). The volatility of Bitcoin and its role as a medium of exchange and a store of value. *Empirical Economics*, *61*, 2663–2683. https://doi.org/10.1007/s00181-020-01990-5

BitInfoCharts. (2025). Top 100 richest Bitcoin addresses. Retrieved March 31, 2025, from https://bitinfocharts.com/top-100-richest-bitcoin-addresses.html

Bleher, J., & Dimpfl, T. (2019). Today I got a million, tomorrow I don't know: On the predictability of cryptocurrencies by means of Google search volume. *International Review of Financial Analysis*, *63*, 147–159. https://doi.org/10.1016/j.irfa.2019.03.003

Bodnaruk, A., & Östberg, P. (2009). Does investor recognition predict returns? *Journal of Financial Economics*, *91*(2), 208–226. https://doi.org/10.1016/j.jfineco.2008.01.006

Bouoiyour, J., Selmi, R., & Tiwari, A. K. (2015). Is Bitcoin business income or speculative foolery? New ideas through an improved frequency domain analysis. *Annals of Financial Economics*, *10*(1), Article 1550002. https://doi.org/10.1142/S2010495215500025

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

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Cheah, E.-T., & Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Conti, M., Kumar, E. S., Lal, C., & Ruj, S. (2018). A survey on security and privacy issues of Bitcoin. *IEEE Communications Surveys & Tutorials*, *20*(4), 3416–3452. https://doi.org/10.1109/COMST.2018.2842460

Cretarola, A., & Figà-Talamanca, G. (2021). Detecting bubbles in Bitcoin price dynamics via market exuberance. *Annals of Operations Research*, *299*(1–2), 459–479. https://doi.org/10.1007/s10479-019-03321-z

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499. https://doi.org/10.1111/j.1540-6261.2011.01679.x

Dastgir, S., Demir, E., Downing, G., Gozgor, G., & Lau, C. K. M. (2019). The causal relationship between Bitcoin attention and Bitcoin returns: Evidence from the Copula-based Granger causality test. *Finance Research Letters*, *28*, 160–164. https://doi.org/10.1016/j.frl.2018.04.019

Dulupçu, M. A., Yiyit, M., & Genç, A. G. (2017). Dijital ekonominin yükselen yüzü: Bitcoin'in değeri ile bilinirliği arasındaki ilişkinin analizi. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, *22*(4), 2241–2256.

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

Foster, F. D., & Viswanathan, S. (1996). Strategic trading when agents forecast the forecasts of others. *The Journal of Finance*, *51*(4), 1437–1478. https://doi.org/10.1111/j.1540-6261.1996.tb04075.x

Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: Feedback cycles between socio-economic signals in the Bitcoin economy. *Journal of the Royal Society Interface*, *11*(99), 20140623. https://doi.org/10.1098/rsif.2014.0623

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C., & Siering, M. (2014). Bitcoin: Asset or currency? Revealing users' hidden intentions. In *Proceedings of the 22nd European Conference on Information Systems (ECIS 2014)*. https://ssrn.com/abstract=2425247

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*(3), 393–408.

Herremans, D., & Low, K. W. (2025). Forecasting Bitcoin volatility spikes from whale transactions and CryptoQuant data using Synthesizer Transformer models. *IEEE Access*, *13*, 117788–117807. https://doi.org/10.1109/ACCESS.2025.3584243

Kapar, B., & Olmo, J. (2021). Analysis of Bitcoin prices using market and sentiment variables. *The World Economy*, *44*(1), 45–63. https://doi.org/10.1111/twec.13020

Kim, G., Shin, D.-H., Choi, J. G., & Lim, S. (2022). A deep learning-based cryptocurrency price prediction model that uses on-chain data. *IEEE Access*, *10*, 56232–56248. https://doi.org/10.1109/ACCESS.2022.3177888

Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In B. E. Eckbo (Ed.), *Handbook of corporate finance: Empirical corporate finance* (Vol. 1, pp. 3–36). Elsevier. https://doi.org/10.1016/B978-0-444-53265-7.50015-9

Kristoufek, L. (2013). Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, *3*,

3415. https://doi.org/10.1038/srep03415

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PLoS ONE*, *10*(4), e0123923. https://doi.org/10.1371/journal.pone.0123923

Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica*, *53*(6), 1315–1335. https://doi.org/10.2307/1913210

Long, S., Xie, Y., Zhou, Z., Lucey, B. M., & Urquhart, A. (2025). From whales to waves: Social media sentiment, volatility, and whales in cryptocurrency markets. *The British Accounting Review*, Article 101682. Advance online publication. https://doi.org/10.1016/j.bar.2025.101682

Mai, F., Shan, Z., Bai, Q., Wang, X., & Chiang, R. H. L. (2018). How does social media impact Bitcoin value? A test of the silent majority hypothesis. *Journal of Management Information Systems*, *35*(1), 19–52. https://doi.org/10.1080/07421222.2018.1440774

Makarov, I., & Schoar, A. (2020). Trading and arbitrage in cryptocurrency markets. *Journal of Financial Economics*, *135*(2), 293–319. https://doi.org/10.1016/j.jfineco.2019.07.001

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Mittal, A., Dhiman, V., Singh, A., & Prakash, C. (2019). Short-term Bitcoin price fluctuation prediction using social media and web search data. In *2019 12th International Conference on Contemporary Computing (IC3)* (pp. 1–6). IEEE. https://doi.org/10.1109/IC3.2019.8844899

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

Nasir, M. A., Huynh, T. L. D., Nguyen, S. P., & Duong, D. (2019). Forecasting cryptocurrency returns and volume using search engines. *Financial Innovation*, *5*(1), Article 2. https://doi.org/10.1186/s40854-018-0119-8

Polasik, M., Piotrowska, A. I., Wisniewski, T. P., Kotkowski, R., & Lightfoot, G. (2015). Price fluctuations and the use of Bitcoin: An empirical inquiry. *International Journal of Electronic Commerce*, *20*(1), 9–49. https://doi.org/10.1080/10864415.2016.1061413

Saggu, A. (2022). The intraday Bitcoin response to Tether minting and burning events: Asymmetry, investor sentiment, and "Whale Alerts" on Twitter. *Finance Research Letters*, *49*, Article 103096. https://doi.org/10.1016/j.frl.2022.103096

Shen, D., & Shi, G. (2025). The role of whale investors in the Bitcoin market. *Research in International Business and Finance*, *78*, Article 103008. https://doi.org/10.1016/j.ribaf.2025.103008

Urquhart, A. (2018). What causes the attention of Bitcoin? *Economics Letters*, *166*, 40–44. https://doi.org/10.1016/j.econlet.2018.02.017

Uzun, M. (2024). The examination of the relationship between Bitcoin (BTC) trading volume in Türkiye and Google Trends data on Bitcoin searches in Google search engine. *Fiscaoeconomia*, *8*(2), 720–738. https://doi.org/10.25295/fsecon.1420143

Yıldırım, Ç. (2020). Google Trends "Bitcoin" aramaları ile Bitcoin/USD fiyatları arasındaki ilişkinin analizi: ARDL sınır testi. *Bilgi Ekonomisi ve Yönetimi Dergisi*, *15*(2), 99–113.

APPENDIX

Table A1. Robustness Check (Forward Cumulative Abnormal Returns, Alternative Thresholds)

Case	Number of Events	Horizon (days)	Mean Event	Mean Control	Mean Diff.	t-stat	p-value
WHALE >100 BTC	734	1	−0.01%	0.00%	−0.01%	−0.09	0.93
WHALE >100 BTC	728	7	0.50%	−0.24%	0.74%	2.01	0.04**
WHALE >100 BTC	724	14	0.92%	−0.39%	1.30%	2.50	0.01**
WHALE >100 BTC	712	28	1.65%	−0.50%	2.14%	2.78	<0.01***
GOOGLE Δ>20%	237	1	0.15%	−0.02%	0.17%	0.55	0.58
GOOGLE Δ>20%	237	7	1.11%	−0.15%	1.26%	1.82	0.07*
GOOGLE Δ>20%	236	14	1.50%	−0.18%	1.68%	1.76	0.08*
GOOGLE Δ>20%	234	28	1.95%	−0.10%	2.04%	1.60	0.11

***, **, * denote significance at the 1%, 5%, and 10% levels, respectively. Mean Difference = Mean(Event) − Mean(Control). Abnormal returns are computed under the constant-mean-return model defined in Section 3 (μ = 0.068% per day). The decline in the number of events as the horizon lengthens reflects data availability, since forward return windows are truncated near the end of the sample.