

Haase, Frederic; Celig, Tom; Rath, Oliver; Schoder, Detlef

Article — Published Version

Wisdom of the crowd signals: Predictive power of social media trading signals for cryptocurrencies

Electronic Markets

Provided in Cooperation with:

Springer Nature

Suggested Citation: Haase, Frederic; Celig, Tom; Rath, Oliver; Schoder, Detlef (2025) : Wisdom of the crowd signals: Predictive power of social media trading signals for cryptocurrencies, Electronic Markets, ISSN 1422-8890, Springer, Berlin, Heidelberg, Vol. 35, Iss. 1, <https://doi.org/10.1007/s12525-025-00815-6>

This Version is available at:

<https://hdl.handle.net/10419/323695>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Wisdom of the crowd signals: Predictive power of social media trading signals for cryptocurrencies

Frederic Haase¹ · Tom Celig¹ · Oliver Rath¹ · Detlef Schoder¹

Received: 19 January 2025 / Accepted: 25 June 2025
© The Author(s) 2025

Abstract

The emergence of cryptocurrencies and decentralized finance (DeFi) applications brings unique challenges, including high volatility, limited fundamental valuation methods, and significant informational reliance on social media. Consequently, traditional trading algorithms and decision support systems (DSS) often fall short in effectively capturing these dynamics, underscoring the need for tailored solutions. Recent research on sentiment analysis in cryptocurrency trading has provided mixed evidence regarding its predictive power, highlighting limitations in generalizability and reliability due to the inherent noise of social media content. Addressing these limitations, this study explores crowd-based trading signals, explicit buy and sell recommendations shared by users on social media platforms including X (formerly Twitter), Reddit, Stocktwits, and Telegram. We apply an event study methodology to analyze over 28,000 trading signals extracted using natural language processing (NLP) techniques based on large language models (LLMs). Our findings demonstrate that these explicit crowd-based signals significantly predict short-term cryptocurrency price movements, particularly for assets with lower market capitalization and recent negative returns. An out-of-sample trading strategy using these signals achieves superior risk-adjusted returns, outperforming both a standard cryptocurrency index (CCI30) and the S&P 500. Additionally, we uncover the role of automated accounts (signal bots) actively disseminating trading recommendations. This research advances literature by introducing a precise alternative to sentiment analysis, contributing to the understanding of social media as a distributed financial information environment, and raising theoretical considerations about algorithmic agency and trust. Practical implications span investors, social media platforms, and regulators.

Keywords Social media signals · Cryptocurrencies · Collective intelligence · Trading signals · Predictive power · Wisdom of crowds

JEL Classification G10 · D8 · D7 · G14 · G41

Introduction

The emergence of blockchain-based cryptocurrencies and decentralized finance (DeFi) applications has fundamentally

transformed digital financial markets. Technologies such as cryptographic tokens, smart contracts, and distributed ledgers have introduced new paradigms for value creation, exchange, and governance (Alt et al., 2024; Beck et al., 2018; Schwiderowski et al., 2024). DeFi innovations have further challenged traditional financial intermediation, introducing novel business models and reducing reliance on centralized institutions (Beinke et al., 2024). Once considered niche innovations, cryptocurrencies have matured into a dynamic asset class, attracting significant interest from both retail and institutional investors (Garcia & Schweitzer, 2015; García-Corral et al., 2022; Phillip et al., 2018). Today, thousands of digital assets are actively traded across centralized and decentralized exchanges, offering features such as decentralized infrastructure, programmable assets, and low transaction costs (Jalal et al., 2021; Li & Whinston, 2020). The approval

Responsible Editor: Christian Janiesch

✉ Frederic Haase
haase@wim.uni-koeln.de
Tom Celig
celig@wim.uni-koeln.de
Oliver Rath
rath@wim.uni-koeln.de
Detlef Schoder
schoder@wim.uni-koeln.de

¹ University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany

of cryptocurrency exchange-traded funds (ETFs) by the US Securities and Exchange Commission (SEC) marks a key milestone toward institutional adoption (Gary, 2024). Yet, the crypto ecosystem continues to face significant challenges, including fragmentation between centralized and decentralized platforms (Hägele, 2024) and regulatory and security concerns (Pocher et al., 2023; Zeißet al., 2024). Despite growing attention, public understanding remains limited, and market behavior is increasingly shaped by the interplay of social and technical forces (Zavolokina et al., 2024).

Cryptocurrency markets differ from traditional asset classes in several aspects: they exhibit high volatility, limited fundamental valuation anchors, and asset values are dominated by influences from social media (Kraaijeveld & De Smedt, 2020; Mai et al., 2018; Piñeiro-Chousa et al., 2023; Tiwari et al., 2025). In this context, sentiment—the aggregated mood, opinions, or emotional tone expressed by users on digital platforms—has emerged as a particularly powerful force, with patterns on platforms such as X (formerly Twitter), Reddit, Telegram, and Stocktwits showing significant correlations with digital asset prices. These platforms are increasingly used for both the dissemination and consumption of financial information (Buz & De Melo, 2024), often shaping perceptions and driving behavior in real-time. Retail investors, particularly younger demographics, are especially active in this space. A recent Forbes Advisor survey reports that 80% of young adults have received financial advice via social media, with a substantial portion of content focused on cryptocurrencies (Egan, 2023). Importantly, institutional investors are also responding to the informational dynamics of social media. Hedge funds, in particular, are not only gaining exposure to digital assets—47% of traditional hedge funds now incorporate cryptocurrencies (pwc & AIMA, 2024)—but are also adapting their trading strategies to the sentiment-driven nature of these markets. Many funds are now incorporating social media sentiment data into algorithmic models and trading systems to scrape platforms like Reddit and Twitter for indicators of sentiment and early signals of price movement (Wilson & Jessop, 2019). In these highly volatile environments, social media sentiment has become a critical input for decision-making.

However, current approaches to extract and use social media sentiment in trading systems face several limitations. Most systems rely on aggregated sentiment indicators derived from large volumes of posts, assuming that averaging will surface dominant market signals (Bari & Agah, 2020). This often introduces substantial noise and weakens predictive power (Subramanian et al., 2023). A major challenge is *linguistic and structural noise*, stemming from the informal and inconsistent nature of social media content. Posts are typically short, ungrammatical, and filled with emojis, tags, and hyperlinks, requiring extensive pre-processing (Kraaijeveld & De Smedt, 2020). Many lack the clarity or

structure needed for traditional natural language processing (NLP) models, leading researchers to employ neural network architectures such as LSTMs or transformers. Yet, accurately capturing sentiment in such unstructured content remains an open challenge (Dong et al., 2024; Subramanian et al., 2023). Equally challenging is *contextual and topical irrelevance*: not every asset mention provides value-relevant information, and many posts reflect background chatter, speculation, or outdated views (Xie et al., 2020). Even on finance-focused platforms like Stocktwits, sentiment can be distorted by varied user intentions and brief, unfocused messages (Deng et al., 2018). While methods like keyword clustering or event co-occurrence have been proposed to improve relevance (Bari & Agah, 2020), the distinction between sentiment and actionable advice remains blurred. A user expressing enthusiasm for an asset may boost positive sentiment scores without offering a clear recommendation to buy, leading to weak signal-to-noise ratios (Subramanian et al., 2023). In short, existing sentiment analysis methods often fail to capture the precision and context required for reliable short-term decision-making in crypto markets, pointing to the need for more targeted alternatives.

To overcome the limitations of sentiment-based approaches, this study focuses on a more direct input: explicit crowd-based trading signals. These are time-stamped buy and sell recommendations shared publicly on platforms like X, Reddit, Stocktwits, and Telegram. Unlike aggregated sentiment, which often reflects general mood or background chatter, these signals represent concrete trading advice and may better capture collective market expectations in real-time. In this study, we examine whether such signals can predict short-term abnormal returns in cryptocurrency markets. Thus, we investigate the following research questions:

- **RQ1:** *Can crowd-based trading signals inform short-term trading decisions for cryptocurrencies and predict abnormal returns?*
- **RQ2:** *For which cryptocurrencies do crowd-based trading signals most effectively inform trading decisions and predict abnormal returns?*

To explore these questions, we extract and analyze 28,700 trading signals from multiple social media platforms using a zero-shot classification approach utilizing large language models (LLMs). We apply an event study methodology to evaluate the predictive power of these signals across 287 cryptocurrencies. Our findings show that crowd-based signals can effectively predict short-term price movements, with the effect size varying based on cryptocurrency-specific characteristics. Our study proposes an exemplary trading strategy based on an out-of-sample dataset and achieves positive risk-adjusted returns compared to several benchmark indices. We

also uncover the presence of automated signal bots that systematically post trading advice.

Theoretically, this study advances the literature on algorithmic trading, financial text mining, and social media's role in cryptocurrency markets in several ways. First, we move beyond traditional sentiment analysis by introducing explicit crowd-based trading signals as an alternative that captures concrete trading intent. Second, we extend the conceptualization of social media from a sentiment proxy to a decentralized financial information infrastructure, where heterogeneous actors generate and diffuse time-sensitive trading advice. Finally, our identification of signal bots raises theoretical questions about algorithmic agency and trust.

Practically, our findings have implications for investors, social media platforms, and regulators. For institutional investors, we show that short-term predictive insights can be extracted from public trading signals utilizing LLMs, supporting their integration as an alternative data source for algorithmic strategies, especially in high-frequency contexts. For retail investors, however, the short time frame in which these signals are predictive may limit their utility and could even be misleading under volatile conditions. For social media platforms, the emergence of trading signals suggests a need for structured mechanisms—such as tagging, filtering, or moderation—to enhance transparency and user experience. Such structuring may also increase the accessibility of these signals for retail users. Lastly, for regulators and policymakers, the rise of automated accounts posting trading advice highlights the need for adapted regulatory frameworks, including disclosure standards, bot detection mechanisms, and social media monitoring tools capable of addressing the unique dynamics of cryptocurrency markets.

The remainder of this paper is structured as follows: the “[Related work](#)” section summarizes related research; the “[Theoretical background and hypotheses development](#)” section outlines the theoretical background and hypotheses; the “[Data collection](#)” and “[Research method](#)” sections present our data and event study methodology; the “[Results](#)” section reports the empirical results; and the “[Discussion](#)” section concludes with implications and directions for future research.

Related work

Predictive power of social media sentiment in financial markets

Investor sentiment, driven by behavioral and psychological factors, has long been recognized as a key influence in financial markets (Kaplanski & Levy, 2010). Negative mood and anxiety, for example, have been shown to affect investment decisions and lead to temporary mispricing. Studies of exoge-

nous shocks, such as aviation disasters, demonstrate that markets can react sharply to sentiment-driven perceptions of risk, with disproportionate short-term losses, even in the absence of fundamental changes. These findings underscore the broader role of sentiment in asset pricing, particularly in environments of uncertainty or limited information (Baker & Wurgler, 2007; Peterson, 2016).

As digital communication platforms became more prominent, social media emerged as a powerful and observable channel through which investor sentiment is expressed. Early work by Dewally (2003) found that internet-based stock recommendations had a modest impact on market returns, suggesting that the informational value of online forums was still developing. However, subsequent research provided a more refined understanding. Antweiler and Frank (2004), for instance, showed that messages on stock message boards significantly correlated with both returns and volatility, highlighting that even low-frequency social media activity could contain meaningful signals for market behavior. Bollen et al. (2011) applied Twitter sentiment analysis to forecast movements in the Dow Jones Industrial Average. Building on these earlier studies, subsequent studies have explored not only the sources of sentiment but also how it should be weighted for predictive purposes. For instance, Chen et al. (2014) found that both news articles and accompanying user comments could predict stock returns and earnings surprises. Similarly, Lachana and Schröder (2025) found that sentiment from social media consistently outperforms traditional media in return prediction. Nofer and Hinz (2015) highlighted that sentiment weighted by follower count improved predictive performance for stocks. The study of Deng et al. (2018) discusses the endogenous relationship between sentiment and financial markets, emphasizing that sentiment not only reflects market conditions but can also actively influence them. The rise of meme stocks serves as a prominent example of how coordinated retail sentiment, amplified through social media, can meaningfully influence market dynamics (Aloosh et al., 2022; Kim et al., 2023). Beyond aggregate sentiment levels, researchers have explored more nuanced dimensions of sentiment derived from social media. See-To and Yang (2017) introduced the concept of sentiment dispersion—variation in opinion—as a driver of increased volatility. Similarly, Li et al. (2018) found that disagreement among posts could predict very short-term price movements.

Following research on social media's influence in stock markets, scholars have increasingly turned to cryptocurrency markets to explore similar dynamics. Given the unique properties of these assets, cryptocurrencies present a unique environment for sentiment-driven trading. Piñeiro-Chousa et al. (2023) found that tweet volume impacts volatility more than returns, suggesting that social media metrics affect perceived risk rather than price levels directly. Xie (2022) analyzed discussions on Bitcointalk.org and noted

that sentiment largely reflected past performance with only limited value-relevant information. Other studies identified more value-relevant information in sentiment: Kraaijeveld and De Smedt (2020) demonstrated the predictive value of Twitter sentiment for major cryptocurrencies, while also noting the distorting influence of bots and noise. Mai et al. (2018) revealed that sentiment from less active users (the *silent majority*) has a stronger correlation with Bitcoin price, highlighting the subtlety of crowd dynamics. Methodological advancements have further refined sentiment-based modeling. Subramanian et al. (2023) incorporated several sentiment dimensions such as emotional tone, factual framing, and informal language into a trading system, resulting in improved prediction accuracy. Meyer et al. (2024) investigated influencers on platforms like YouTube and their association with Bitcoin prices, discussing the aspect of source credibility. Garcia and Schweitzer (2015) observed that shifts in Twitter opinion polarization precede Bitcoin price movements.

To contextualize our contribution within this evolving literature, Table 1 summarizes key studies across cryptocurrency sentiment research, outlining differences in platform focus, analytical methods, and temporal resolution. While most prior work relies on aggregated sentiment indicators, we take a different approach by extracting explicit buy and sell signals from multiple platforms using a zero-shot

LLM framework. We then apply an hourly event study methodology to a broad set of cryptocurrencies, enabling a high-resolution analysis of short-term market responses to public trading signals. Our choice of hourly data is motivated by the need to accurately capture high-frequency intraday dynamics in cryptocurrency markets. Prior studies using daily data are limited in this aspect to reflect price fluctuations that can occur within a single day (Xie, 2022). We also extend the analysis to multiple cryptocurrencies, given that prior empirical findings highlight substantial heterogeneity across cryptocurrencies in their associations with financial markets (Kraaijeveld & De Smedt, 2020).

Noise in social media sentiment of financial markets

Common practice in utilizing social media data is to aggregate sentiment scores from large numbers of social media posts, with the goal of extracting actionable insights (Bari & Agah, 2020). This approach assumes that averaging will filter out irrelevant or misleading content, allowing the dominant signal to surface. However, this method often fails to capture fine-grained or context-specific sentiment shifts and instead incorporates significant noise, weakening its usefulness in trading strategies (Subramanian et al., 2023).

One major issue is *linguistic and structural noise*, stemming from the informal, unstructured nature of social media

Table 1 Comparison of related work on social media and cryptocurrency markets

Study	Time Res.	Assets	Social media source	Signal type	Method/model
Piñeiro-Chousa et al. (2023)	Weekly	51 DeFi products	Twitter	Tweet metrics	Panel regression
Xie (2022)	Hourly	Bitcoin	Bitcointalk.org	Discussion sentiment	VARX regression
Kraaijeveld and De Smedt (2020)	Hourly, Daily	9 cryptocurrencies	Twitter	Discussion sentiment	Granger causality
Mai et al. (2018)	Daily	Bitcoin	Forum, Twitter	User subgroup sentiment	VECM regression
Subramanian et al. (2023)	Weekly	Bitcoin	News, social media (MarketPsych)	Sentiment indices	LSTM neural network
Meyer et al. (2024)	Hourly	Bitcoin	YouTube	Influencer sentiment	Event study
Xie et al. (2020)	Daily	Bitcoin	Bitcointalk.org	Weighted sentiment	Panel regression
Garcia and Schweitzer (2015)	Daily	Bitcoin	Twitter	Emotional valence & opinion polarization	VAR regression
This study	Hourly	287 cryptocurrencies	X, Reddit, Stocktwits, Telegram, Discord	Direct trading signals	Event study

posts—especially within financial subcommunities, where domain-specific slang and abbreviations further complicate interpretation (Agrawal et al., 2022). As shown by Kraaijeveld and De Smedt (2020), extensive pre-processing is required to clean this data, and many posts are excluded entirely due to the lack of useful content. While hashtags can provide some semantic cues, they often require additional interpretation to be meaningful. This type of noise is particularly problematic for NLP models that attempt to extract sentiment without accounting for the quality or coherence of the language used. Subramanian et al. (2023) emphasize that traditional sentiment extraction techniques struggle with this issue, and propose neural filtering—such as LSTM-based models—as a way to differentiate signal from syntactic or linguistic noise. Likewise, Dong et al. (2024) show that low-quality posts significantly hinder prediction accuracy unless filtered through more advanced systems designed to assess textual quality. Yet, even these advanced methods continue to face challenges in effectively handling the noise present in financial social media data.

The second key challenge is *contextual and topical irrelevance*. Not every post that mentions an asset contributes meaningful or timely information. Many are speculative, repetitive, or disconnected from actual market developments. Even platforms like Stocktwits, which focus more explicitly on financial discourse and use cashtags to map posts to specific assets, are still subject to noise from varying user intentions and the brevity of messages (Deng et al., 2018). In many cases, aggregated sentiment reflects background chatter rather than focused, event-driven discourse. Xie et al. (2020) argue that the structure and cohesion of discussion networks are critical: sentiment signals from loosely connected conversations are less predictive and can even introduce spurious relationships. Bari and Agah (2020) propose that clustering tweets by keyword similarity and event co-occurrence—rather than relying on raw volume can better capture relevance and filter out extraneous content. This suggests that identifying value-relevant signals requires not only textual filtering but also structural and topical alignment across messages.

To address these challenges, our approach focuses on filtering and analyzing explicit trading signals. This approach helps mitigate both linguistic noise and contextual irrelevance, offering a more actionable foundation for short-term trading analysis and incorporation into trading systems.

Previous research on crowd-based trading signals

Several research streams explore crowd-based trading signals, as understood in our study, but in different contexts. One stream has investigated how crowd-based trading signals are used on social media platforms to coordinate illegal pump-and-dump activities. These schemes, characterized by

coordinated efforts to artificially increase the price of a cryptocurrency before selling it, pose significant challenges to market integrity (Kamps & Kleinberg, 2018; La Morgia et al., 2023). In these schemes, fraudsters publish buy and sell signals on social media to manipulate asset prices. La Morgia et al. (2023) conducted a comprehensive analysis of pump-and-dump schemes based on signals, particularly those organized by online communities on Telegram and Discord. In addition, researchers investigated target-based pump signals that publish buy recommendations together with target prices on social media (Hamrick et al., 2021). It was found that half of these target-based pump signals succeeded in driving the cryptocurrency's price to the specified target post-release. Another study proposed a neural network based approach to predict the target cryptocurrency of a pump scheme before its signal release, leveraging market and social media data (Nghiem et al., 2021). Moreover, a related stream of research has examined the impact of influential figures such as Donald Trump (Gjerstad et al., 2021), Elon Musk (Ante, 2023), or social media influencers (Meyer et al., 2024; Haase et al., 2025).

Another stream of literature analyzed crowd-based trading signals from copy trading platforms. These platforms are specifically designed to enable participants to access information about the successes of other traders and replicate their trading strategies (Apesteguia et al., 2020). Several researchers have investigated signals from copy trading platforms for their predictive power. For example, Breitmayer et al. (2019) focused on the social investment platform Sharewise, finding that collective stock assessments could predict stock performance, yielding a monthly excess return of 3.3%. Likewise, research on eToro, an exchange copy trading platform, highlighted the influence of social networks in trading decisions (Pan et al., 2012). Trades influenced by social connections typically outperformed individual trades, although the social reputation of the top traders was not correlated with their performance.

Our investigation closely relates to the study conducted by Buz and De Melo (2024), investigating signals posted on Reddit's Wallstreetbets (WSB) for stocks. The findings indicate that these signals outperform leading investment banks at detecting top-performing stocks. Their research shows that WSBs buy signals for S&P 500 stocks achieve about 70% accuracy over three months, comparable to top investment banks, with a notable price increase of 7% when utilizing specific filtering criteria of signals. These results demonstrate the potential predictive value of crowd-sourced trading signals in the stock market. Complementing these equity-market insights, recent work in decentralized finance shows that follower count on a DeFi portfolio-sharing platform is positively associated with subsequent portfolio returns (Celig et al., 2024).

Building on these research streams, our study examines crowd-based trading signals derived from general-purpose social media platforms. Our study moves beyond pump-and-dump schemes, influencer-driven effects, and copy trading platforms by analyzing diverse, user-generated signals across multiple platforms. Unlike Buz and De Melo (2024), who focus on long-term stock predictions on Reddit, our study investigates post-level, short-term effects in cryptocurrency markets.

Theoretical background and hypotheses development

This section outlines the theoretical foundations and rationale guiding our study. We begin by introducing signaling theory as a lens to understand the interaction between actors issuing cryptocurrency trading signals and those interpreting them. We then discuss the wisdom of the crowd theory, which underpins the collective predictive power attributed to aggregated social media signals. Drawing on these perspectives, we develop our hypotheses.

Signaling theory and the signaling crowd

We conceptualize the emergence of *signals* and the *signaling crowd* in cryptocurrency markets through the lens of signaling theory. In our context, signals refer to explicit trading recommendations disseminated on social media, urging immediate buy or sell actions for specific cryptocurrencies. These signals differ from general news, discussions, entertainment content, or long-term investment advice commonly found on social platforms (Buz & De Melo, 2024). The emphasis on “immediate” reflects the urgency conveyed by directives such as “buy now” or “sell now.” The signaling crowd encompasses a wide range of actors producing these actionable signals, including automated trading bots, financial influencers (finfluencers), copy trading communities, professional analysts, legitimate trading platforms, market manipulators involved in pump-and-dump schemes, and individual investors.

Signaling theory (Connelly et al., 2011; Spence, 2002) provides a useful framework for analyzing the interaction between signalers, those who issue trading signals, and receivers (investors and traders) interpreting them. Unlike traditional market insiders with privileged information (Yasar et al., 2020), social media signalers vary widely in credibility, motivation, and access to information. These signals are publicly observable and widely accessible, and the costs of signaling are often reputational rather than financial (Talmor, 1981). For receivers, the challenge might not be a lack of information, but rather an overabundance of it, given the volume of signals and discussions across platforms.

This highlights the importance of algorithmic consideration of such signals for decision-making. In what follows, we apply this signaling-theoretic lens to characterize the phenomenon of trading signals and their role in shaping behavior within cryptocurrency markets.

Wisdom of the crowd

The *wisdom of the crowd* theory suggests that aggregated judgments from a diverse group often outperform those of individual experts in terms of accuracy and decision quality (Surowiecki, 2005), making it a valuable foundation for analyzing crowd-based predictive signals. Empirical evidence from various domains supports this idea, showing that collective judgments can match or exceed expert evaluations (Wagner & Vinaimont, 2010). Financial markets offer a particularly relevant context. For example, Nofer and Hinz (2014) found that stock predictions from online communities outperformed those of professional analysts, highlighting the value of aggregated crowd recommendations. Similarly, Gottschlich and Hinz (2014) showed that stock votes from online investment communities could be transformed into effective investment strategies using decision support systems, outperforming market benchmarks. Chen et al. (2019) emphasized the role of monetary incentives in shaping the quality of crowd-sourced stock opinions on social media, illustrating how incentive structures affect collective decision-making. More recent advances, such as artificial prediction markets proposed by Dong et al. (2024), offer mechanisms to dynamically distill the wisdom of the crowds from noisy financial online discussions. However, the effectiveness of crowd wisdom depends on specific conditions: diversity, independence, and decentralized aggregation of opinions (Larrick et al., 2012). We posit that these conditions are likely met in broad social media environments, where a wide range of actors independently share trading signals.

Hypotheses development

Drawing on signaling theory and the wisdom of crowds lens, we theorize that individual trading signals shared on social media can serve as meaningful reflections of market expectations. When such signals are issued across a large and diverse group of users, their aggregation can produce what prior research has described as collective intelligence: the idea that crowds, under the right conditions, can outperform individuals or experts in making predictions (Surowiecki, 2005; Nofer & Hinz, 2014; Gottschlich & Hinz, 2014). However, the accuracy of this collective judgment depends on the ability to identify and isolate high-quality signals. Prior approaches, such as sentiment aggregation, often struggle with this due to the noisy nature of social media content.

Our approach addresses this by focusing on direct trading signals. Such signals are more likely to reflect intentional, decision-relevant information. Based on this reasoning, we hypothesize the following:

H1a: *Buy signals predict significant positive returns in cryptocurrencies.*

H1b: *Sell signals predict significant negative returns in cryptocurrencies.*

Prior research shows that the strength of the relationship between social media sentiment and financial markets varies across cryptocurrency assets (Kraaijeveld & De Smedt, 2020). Building on this insight, we hypothesize that specific attributes of a cryptocurrency influence also how strongly buy or sell signals are associated with short-term price movements. This perspective informs our further hypothesis development by motivating the inclusion of these asset-specific attributes.

Effect of cryptocurrency market capitalization

The EMH posits that asset prices in efficient markets fully reflect all available information (Fama, 1965). However, in cryptocurrency markets, efficiency varies considerably across assets and is often linked to characteristics such as market capitalization. Prior research offers mixed evidence on this relationship. Sigaki et al. (2019) report widespread inefficiencies without a consistent connection to market capitalization, implying that both small and large cryptocurrencies can exhibit informational inefficiencies. In contrast, Kang et al. (2022) find that cryptocurrencies with a higher market capitalization tend to be more efficient, suggesting lower information asymmetry and more accurate price discovery.

These findings align with established results from traditional financial markets. Larger firms are generally associated with greater market efficiency, while smaller firms are more prone to inefficiencies and behavioral anomalies (Yang & Pangastuti, 2016). Chung and Hrazdil (2010) further show that liquidity—which often correlates with market capitalization—enhances efficiency.

Building on these mechanisms, we propose that cryptocurrencies with smaller market capitalization—due to their lower efficiency—are more likely to exhibit predictable short-term price patterns in response to crowd-based trading signals. Therefore, we hypothesize:

H2a: *Buy signals are more predictive of positive returns for low market capitalization cryptocurrencies.*

H2b: *Sell signals are more predictive of negative returns for low market capitalization cryptocurrencies.*

Effect of cryptocurrency performance

Cryptocurrency performance, commonly measured by average returns, reflects recent financial success or decline in an asset. Variations in performance have been linked to factors such as market attention, innovation potential, and speculative interest (Wang & Vergne, 2017). Prior studies show that markets tend to be more efficient during periods of strong performance, and more prone to inefficiencies during downturns (Zhang et al., 2020). Levich et al. (2019) find that excess return predictability—an indicator of inefficiency—is particularly pronounced during unstable or underperforming market conditions. Similarly, Yaya et al. (2021) report that inefficiencies and volatility persistence in cryptocurrency markets intensify during periods of distress or weak performance, suggesting that signal predictability may depend on broader market conditions.

We argue that during periods of poor performance, cryptocurrencies are more susceptible to behavioral trading, and pricing inefficiencies. These are conditions under which crowd-based trading signals may be especially effective. Based on this reasoning, we propose:

H3a: *Buy signals are more predictive of positive returns for cryptocurrencies in times of low performance.*

H3b: *Sell signals are more predictive of negative returns for cryptocurrencies in times of low performance.*

Effect of cryptocurrency age

The age of a cryptocurrency affects its credibility, investor trust, and market efficiency. Established cryptocurrencies such as Bitcoin or Ethereum generally exhibit higher liquidity, greater visibility, and more stable trading behavior. These factors contribute to more efficient markets where prices rapidly incorporate available information (Kang et al., 2022; Wei, 2018). In contrast, newer and less-established cryptocurrencies tend to have lower liquidity and visibility, and exhibit greater price volatility and inefficiencies (Wei, 2018).

This efficiency gap is central from the perspective of signaling theory, which emphasizes that signals are particularly valuable in settings with high informational asymmetries (Spence, 1973). In efficient markets—typically seen with older cryptocurrencies—informational asymmetries are reduced, limiting the incremental value of crowd-based signals. However, in less efficient markets, such as those of newer cryptocurrencies, persistent asymmetries create greater potential for signals to convey information not yet reflected in prices.

Based on this theoretical foundation and prior empirical findings, we expect crowd-based signals to be more

informative for newer cryptocurrencies. Accordingly, we hypothesize the following:

H4a: *Buy signals are more predictive of positive returns for non-established cryptocurrencies.*

H4b: *Sell signals are more predictive of negative returns for non-established cryptocurrencies.*

Data collection

For our research, we collected two datasets (i) crowd-based trading signals for several cryptocurrencies and (ii) their corresponding market prices and metadata such as market capitalization. As the generation of the crowd-based trading signal dataset is seminal to our study, we detail the data-collection methodology behind this dataset in the following subsection.

Crowd-based trading signals

In our research, we compiled a dataset of crowd-based trading signals by collaborating with *Stockpulse*, a provider with extensive history of financial social media data. The database of *Stockpulse* granted us access to a wide range of financial social media data from major platforms including X, Stocktwits, Telegram, Reddit, and more. This comprehensive coverage allowed us to collect content, such as posts or comments, from various platforms. We utilized a programmatic search interface to extract data for individual cryptocurrencies based on keyword searches. For this study, we selected a time frame from January 1, 2022, to June 30, 2023, to encompass a diversity of market periods and market events (such as the FTX exchange collapse (Bouri et al., 2023)) across the cryptocurrencies. We queried data for the 9000 cryptocurrencies listed on *CoinMarketCap*.¹ To generate a substantial volume of potential crowd-based trading signals, we conducted the following steps:

1. **Search for signals:** Our approach to extract crowd-based trading signals begins with identifying social media posts that reference individual cryptocurrencies and contain potential trading advice. Specifically, we selected posts that (i) include exactly one cashtag (e.g., \$BTC for Bitcoin), allowing unambiguous association with a single cryptocurrency, and (ii) contain the keyword *signal*, which we used as a minimal criterion to identify posts likely related to trading advice. The list of valid cashtags was based on the approximately 9000 cryptocurrencies listed on *CoinMarketCap* during our study period. For each

post, we combined the title (if available) and body into a single input text, removed entries with missing or excessively long content, and filtered for cashtags of assets. This process yielded a dataset of approximately 480,610 candidate posts covering 2496 different cryptocurrencies. These posts formed the input for the subsequent classification step.

2. **Zero-shot classification of buy and sell recommendations:** To automatically categorize the extracted posts into actionable trading signals, we applied a zero-shot classification approach using the *facebook/bart-large-mnli* LLM (Lewis et al., 2020; Williams et al., 2018; Yin et al., 2019). Zero-shot classification is a technique that allows a model to assign labels to input text without having seen labeled examples for the specific task during training. Instead, the model evaluates the semantic similarity between a given text and a set of candidate labels expressed in natural language. This method is a form of transfer learning, enabling general-purpose language models to perform downstream classification tasks with minimal task-specific engineering (Yin et al., 2019). Zero-shot, single-shot, and few-shot classifications are considered emergent capabilities of LLMs, typically becoming effective at large model sizes. The BART-large model, with over 400 million parameters, is well-suited for this task. In our context, zero-shot classification allows us to assign directional trading labels at scale without relying on a large annotated dataset, providing a practical alternative to rigid keyword-based methods.

Each candidate social media post was assessed against five candidate labels: *Immediate buy advice*, *Immediate sell advice*, *Neutral advice*, *Long-term trading advice*, and *No trading advice*. The label with the highest confidence score was assigned to the post. The first three categories—buy, sell, and neutral—are inspired by Geritsen et al. (2022), who proposed a bearish/neutral/bullish scale to differentiate cryptocurrency analysts' sentiment. We added the label *Long-term trading advice* to explicitly capture investment-oriented content that is inconsistent with our focus on short-term effects. Finally, the label *No trading advice* was included to exclude posts unrelated to trading or financial advice. The distribution of classified posts was as follows: 53.4% were labeled as *Immediate buy advice*, 27.0% as *Immediate sell advice*, 11.6% as *Neutral advice*, 6.9% as *Long-term trading advice*, and 1.1% as *No trading advice*.

To ensure the quality of the automatically extracted signals, we conducted a manual validation study. A random sample of 200 classified posts (100 buy and 100 sell) was independently reviewed by three human annotators and checked whether it correctly contained immediate trading advice. The accuracy of the zero-shot classification was 80% for buy signals and 82% for sell signals. To mea-

¹ <https://coinmarketcap.com>.

sure inter-rater agreement, we computed Gwet's Gamma (Gwet, 2014), which is particularly robust in the presence of class imbalance, given the high accuracy. Agreement scores were substantial for both categories, with Gwet's Gamma of 0.7239 for buy signals and 0.7179 for sell signals. We concluded that our signal extraction approach has sufficient quality and we retained only the posts classified as either *Immediate buy advice* or *Immediate sell advice* for further analysis, resulting in a dataset of 256,650 buy signals and 129,649 sell signals.

3. **Balancing dataset:** In the resulting dataset, we noted an imbalance in the number of signals associated with each cryptocurrency. A significant number of the signals were from the cryptocurrencies Bitcoin and Ethereum. To address this imbalance and ensure a more uniform dataset, we established a criterion for inclusion based on the number of signals per cryptocurrency. Specifically, we included only those cryptocurrencies that had a minimum of 100 signals (either buy or sell) during our study period. This threshold was chosen based on the median number of signals per cryptocurrency and preliminary analyses that indicated that this threshold provides a substantial volume of signals for statistical analysis. For cryptocurrencies that exceeded this threshold, due to an abundance of signals, we randomly selected 100 signals to maintain uniformity across our dataset. The balancing serves two purposes. First, it prevents our analysis from being biased by a few cryptocurrencies with many signals. Second, we can effectively analyze the effects of various cryptocurrency properties and make the signals comparable across different cryptocurrencies. This process resulted in a final dataset including 287 cryptocurrencies (see Appendix A), each with 100 signals. Our final dataset contains $100 \times 287 = 28,700$ signals.

Table 2 displays the total number of signals listed for each source. We observed that most of the signals were extracted from the social media platform X, followed by Reddit, Stock-

twits, and Telegram. Figure 1 offers an hourly aggregation of the count of crowd-posted signals. Table 3 provides signal examples. Given the prevalence of pump-and-dump schemes in cryptocurrencies, we compared our set of cryptocurrencies with one of the largest public datasets of such schemes (La Morgia et al., 2023). We found that approximately 10% of the cryptocurrencies in our dataset have been historically targeted by pump-and-dump schemes, however, no known pump-and-dump attempt is matched to a signal in our dataset. Although we therefore cannot completely rule out the possibility that our dataset contains pump-and-dump attempts, we included signals from these cryptocurrencies in our analysis as we consider them as part of the overall signal phenomenon.

Cryptocurrency data

The price data for the 287 cryptocurrencies in scope of this study was retrieved from CoinMarketCap, which has been used in prior research (e.g., Momtaz (2021)). CoinMarketCap provides aggregated pricing data from reputable and leading exchanges such as Coinbase, Kraken, and others and covers the majority of cryptocurrencies. The price data was fetched in hourly resolution for the whole time frame of January 1, 2022, to June 30, 2023. We selected hourly end data as aggregation for each hour. Additionally, we collected data about the market capitalization and ages of cryptocurrencies on CoinMarketCap.

Research method

The event study method provides a framework for analyzing the immediate predictive power of signals on cryptocurrency returns. It is a widely recognized method in financial and IS research. This method has been used in other contexts for analyzing the relation between events and the cryptocurrency markets (e.g., Shanaev et al. (2019) and Yue et al. (2021)). Henderson (1990) highlights the versatility and effectiveness of event studies, even under suboptimal conditions, making them a suitable tool for the dynamic market of cryptocurrencies. An in-depth description of the method described below can be found in various research (Dos Santos et al., 1993; Im et al., 2001; Bose & Pal, 2012). We conducted our event study in hourly resolution to investigate short-term effects of signals. However, event studies do not necessarily indicate whether signals will be profitable in an actual trading strategy. Therefore, we additionally conducted an analysis of a hypothetical trading strategy to determine if these signals could be used profitably. While it is not advised to purely build a trading strategy based on these signals alone, we aim to isolate and examine their predictive power. It is impor-

Table 2 Number of collected signals by source. The majority of signals are collected from X, highlighting that a significant portion of the signals are published on this platform

Source	Buy signals	Sell signals
X	17472	10071
Reddit	355	114
Stocktwits	263	126
Telegram	115	82
CoinMarketCap	38	13
Discord	20	31

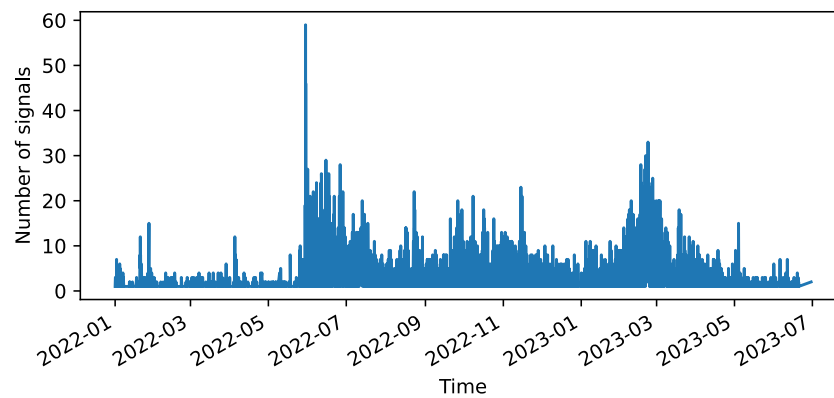


Fig. 1 Hourly number of signals

tant to note that it is not the core contribution of the paper to design a profitable trading strategy. Rather, it serves as an auxiliary analysis to complement our event study findings.

Event study

For the computation of abnormal returns, Fama and French (2004) capital asset pricing model (CAPM) was used as shown in Formula 1:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \quad (1)$$

where R_{it} represents the rate of return for signal i in hour t , while R_{mt} denotes the rate of return for the cryptocurrency market index m , as detailed in the “[Cryptocurrency market](#)

[index](#)” section, in hour t . The y-intercept α and the slope β reflect the sensitivity to R_{mt} . ε is the error term. We utilize 240 h of data before each signal for fitting the regression model. Additionally, a 24-h gap is maintained between the estimation window and the event window, to mitigate the event’s influence on the estimation model. We conducted a robustness check using a market-adjusted model and discuss the selection of CAPM in a later section. The formula to determine the abnormal return (AR) for signal i on hour t is outlined in 2:

$$AR_{it} = R_{it} - (a_i + b_i R_{mt}) \quad (2)$$

where R_{it} represents the rate of return, and R_{mt} denotes the cryptocurrency market index, while a_i is the intercept and

Table 3 Sample of collected signals

Datetime	Text	Ticker	Signal	Source
2023-03-18 16:05:00	#BUY \$AVAX! +2.315% in 10 seconds Current price: 18.374\$! Follow for more signals!	\$AVAX	BUY	X
2022-07-12 14:30:00	\$NNT flashing a clean MACD buy signal right now. Worth watching. [...]	\$NNT	BUY	Telegram
2023-05-08 20:00:00	\$MINA is now printing a buy signal on the daily chart. Potential break out inbound. This aged incredibly well	\$MINA	BUY	Reddit
2022-01-06 17:00:00	\$MANA Buy signals on the 30 min and 1 hour	\$MANA	BUY	Stocktwits
2023-06-12 09:45:00	[...] New short signal detected for \$SHIB. [...]	\$SHIB	SELL	X
2023-02-17 20:00:00	\$FIL got a signal - strong sell now!	\$FIL	SELL	CoinMarketCap
2023-01-27 11:15:00	#SELL signal triggered for \$MATIC! -1.785% in 15min. Monitor closely. [...]	\$MATIC	SELL	Reddit
2023-03-15 12:00:00	Big short coming soon! Sell now signal \$MASK	\$MASK	SELL	Telegram

b_i is the slope of the regression model corresponding to the signal i . The calculation of the abnormal return (AR) is subsequently derived based on Formula 3:

$$AR = \frac{\sum_{i=1}^N AR_{it}}{N} \quad (3)$$

where N describes the total number of signals. To perform statistical significance tests, the abnormal returns (AR_{it}) are converted into the standardized abnormal return (SAR_{it}), as shown in Formula 4:

$$SAR_{it} = \frac{AR_{it}}{\sqrt{\text{Var}(AR_{it})}} \quad (4)$$

In this formula, $\text{Var}(AR_{it})$ is defined as $s_i^2 \left[1 + \frac{1}{D_i} + \frac{(R_{mt} - R_m)^2}{\sum_{j=-T_1}^{-T_2} (R_{mj} - R_m)^2} \right]$, where s_i^2 represents the variance of residual returns from the regression model for signal i . D_i denotes the count of hours used in the regression for signal i , extending from T_1 to T_2 hours before the signal's event hour. R_m stands for the average return on the cryptocurrency market index over the estimation span of D_i . The standardized abnormal return (SAR) for all signals is then calculated as per Formula 5:

$$SAR = \frac{1}{N} \sum_{i=1}^N \frac{SAR_{it}}{\sqrt{2}}. \quad (5)$$

To evaluate the statistical significance of the abnormal returns, we applied two non-parametric testing procedures (Anh & Wolf, 2023; Kolari & Pynnonen, 2011). These methods were chosen because they are more robust than parametric alternatives to violations of normality and the presence of outliers, features that are common in cryptocurrency return data. The first approach is a generalized rank test, which compares the distribution of event-window ranks to those observed in the estimation window. Let R_1 be the sum of ranks for the event sample of size n_1 , and let n_2 be the number of observations in the estimation sample. The expected rank sum and its standard deviation are computed under the null hypothesis as $E(R_1) = \frac{n_1(n_1+n_2+1)}{2}$ and $\sigma(R_1) = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$. The corresponding test statistic is defined as:

$$Z_{\text{rank}} = \frac{R_1 - E(R_1)}{\sigma(R_1)} \quad (6)$$

and the associated two-tailed p -value is derived from the standard normal distribution.

Additionally, we run a permutation test to assess the significance of the observed AR . This approach generates

a null distribution by randomly sampling values from the estimation period, thereby avoiding assumptions about the underlying distribution. Let S_{obs} be the observed average AR during the event window, and S_b denote the AR of the b -th random sample drawn from the estimation set. Repeating this process B times yields the following test statistic:

$$p_{\text{perm}} = \frac{1}{B} \sum_{b=1}^B \mathbb{I}(|S_b| \geq |S_{\text{obs}}|) \quad (7)$$

where \mathbb{I} is the indicator function. We set $B = 100$.

Cryptocurrency market index

We develop a capitalization-weighted index due to the lack of an established cryptocurrency market index in hourly resolution. We follow the index construction described by Stotz et al. (2010). The index was calculated with an hourly resolution, including all 287 cryptocurrencies in scope of this study. This index is crucial for calculating the rate of return, R_{mt} as defined by the CAPM model. Given the large market capitalization of Bitcoin and Ethereum, there is a noticeable correlation of the index with these cryptocurrencies. The formula for the capitalization-weighted index is presented as follows:

$$I_t = \frac{\sum_{c=1}^n (P_{ct} \cdot M_{ct})}{\sum_{c=1}^n M_{ct}} \quad (8)$$

In this formula, P_{ct} is the price of a given cryptocurrency c at hour t , M_{ct} represents the market capitalization of that cryptocurrency at a given hour and n signifies the total number of cryptocurrencies included in the index. Figure 2 illustrates the index price. To calculate the rate of return R_{mt} for the market index at hour t , we use the following formula:

$$R_{mt} = \frac{I_t - I_{t-1}}{I_{t-1}} \quad (9)$$

Subsample analysis

To test hypotheses H2–H4, we used both non-parametric tests for significance. First, we created subsamples of signals based on each hypothesis's criteria and calculated their abnormal returns. We then compared these returns across different subsamples of signals. For subsampling, cryptocurrencies were categorized based on factors relevant to hypotheses H2–H4. Specifically, cryptocurrencies were ranked by market capitalization as of June 30, 2023, and divided at the median into two groups: one comprising those with lower market capitalization and the other with higher capitalization, resulting in two equal-sized signal subsamples. For the performance analysis, cryptocurrencies

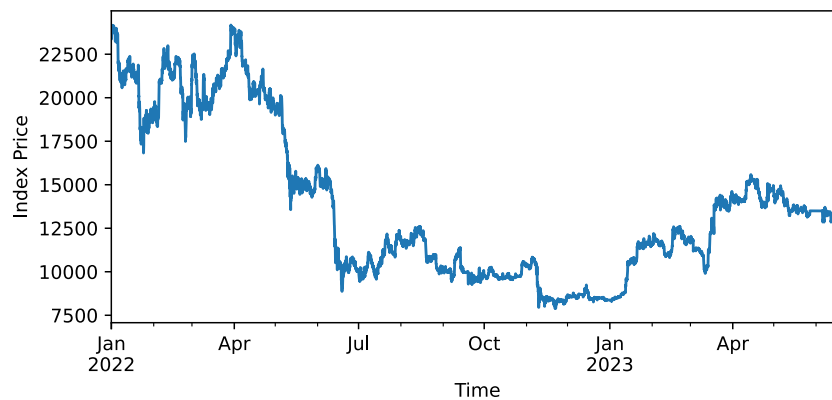


Fig. 2 Market capitalization weighted index price over time

were classified according to their average returns during the study period, with the median used to separate high-return from low-return cryptocurrencies. Lastly, for analyzing cryptocurrency age, the median age of the 189 cryptocurrencies included in the study, as recorded on CoinMarketCap as of June 30, 2023, was used to divide them into older and newer groups. Table 4 details the properties defining these subsamples.

Finally, we performed a Pearson correlation analysis between all factors used in subsampling, reported in Table 5. Each cryptocurrency was assigned a binary value (0 or 1) corresponding to its subsampling factor. For instance, cryptocurrencies with high market capitalization were assigned a value of 1, while those with small market capitalization were

assigned 0. We then computed pairwise correlations between the three subsampling factors. The correlations ranged from -0.302 to 0.129 , indicating only weak relationships among these factors. This suggests that each factor can be analyzed independently in subsequent hypothesis testing.

Results

Event study

Tables 6 and 7 report abnormal returns (AR) around cryptocurrency buy and sell signals, analyzed hourly from 8 h before to 8 h after signal release. Similar hourly time spans were investigated for example by Xie (2022). Specifically, we examined AR up to 8 h before the signal hour to detect any potential leakage, at the signal hour (i.e., 0 h) to assess immediate predictive power, and up to 8 h post signals to evaluate delayed effects.

For buy signals, we observe the most significant effect at event hour 0, with a statistically significant positive abnormal return of 0.003497 (permutation test $p < 0.01$), and 55.83% of signals resulting in positive returns. This supports the hypothesis (H1a) that buy signals are statistically associated with immediate positive returns. Preceding the signal, we find evidence of potential information leakage or reactive signal issuance: positive AR values at hours -6 and -2 are statistically significant with p -values below 0.05, despite smaller magnitudes (e.g., $AR = 0.000647$ at hour -6). Post-

Table 4 Determination of subsamples

Panel A: Cryptocurrency market capitalization	
Statistic	Market capitalization in USD
Mean	116,371,682
Standard deviation	5,768,235,776
25%	55,690,677
50%	184,942,035
75%	569,739,324
Panel B: Cryptocurrency performance	
Statistic	Average return
Mean	-0.000102
Standard deviation	0.000225
25%	-0.000131
50%	-0.000085
75%	-0.000048
Panel C: Cryptocurrency age	
Statistic	Coin age
Minimum	2010-07-13
Maximum	2022-01-31
Median	2019-10-25

Table 5 Correlation analysis of subsamples

	(1)	(2)	(3)
(1) Market capitalization	1	0.129	-0.108
(2) Performance	0.129	1	-0.302
(3) Age	-0.108	-0.302	1

Table 6 Event study results for buy signals using generalized rank and permutation tests. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Hour	AR	Pos. returns	Generalized rank test (Z_{rank})	Permutation test (p_{perm})
−8	0.000457	52.60%	0.080*	0.110
−7	0.000243	52.03%	0.101	0.180
−6	0.000647	52.43%	0.060*	0.020**
−5	0.000195	52.54%	0.080*	0.220
−4	0.000292	51.50%	0.554	0.100*
−3	0.000440	51.58%	0.128	0.090*
−2	0.000424	51.89%	0.020**	0.030**
−1	−0.000225	51.29%	0.964	0.090*
0	0.003497	55.83%	0.000***	0.000***
1	−0.000469	50.51%	0.028**	0.000***
2	0.000102	51.55%	0.628	0.510
3	−0.000075	51.22%	0.881	0.530
4	−0.000135	51.31%	0.951	0.320
5	−0.000120	50.93%	0.093*	0.330
6	−0.000093	50.41%	0.069*	0.430
7	−0.000198	50.43%	0.020**	0.170
8	−0.000029	51.35%	0.643	0.780

event, however, results become ambiguous. Hour +1 shows a statistically significant negative AR of −0.000469, possibly indicating price reversals or overreaction. Beyond this point, no clear statistically significant effects are found. Overall, these results support H1a, particularly for short-term predictive power of buy signals.

For sell signals, event hour 0 again yields the strongest effect, with a statistically significant negative AR of −0.005180 ($p < 0.01$), and 63.87% of signals leading to negative

returns. This confirms H1b, pointing to strong immediate predictive power of sell signals. Unlike the buy case, the largest pre-event effect appears at hour −1 ($AR = 0.001693$), which is significant across both tests, suggesting a clear pattern of signals issuance following price increases. After the signal, the effects are less consistent. However, hour +4 receives a significant negative AR (−0.000530, $p < 0.01$), and similar significance is observed at hours +6 and +7, suggesting that sell signals may have sustained predictive value. Thus, we

Table 7 Event study results for sell signals using generalized rank and permutation tests. Significance levels: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Hour	AR	Neg. returns	Generalized rank test (Z_{rank})	Permutation test (p_{perm})
−8	−0.000468	51.47%	0.101	0.044**
−7	0.000378	48.19%	0.172	0.070*
−6	−0.000075	50.58%	0.091*	0.770
−5	−0.000137	50.62%	0.040**	0.380
−4	0.000079	48.84%	0.067*	0.700
−3	−0.000065	50.02%	0.147	0.710
−2	0.000448	48.94%	0.086*	0.030**
−1	0.001693	45.41%	0.000***	0.000***
0	−0.005180	63.87%	0.000***	0.000***
1	0.000395	49.77%	0.414	0.060*
2	0.000214	49.12%	0.701	0.280
3	−0.000171	50.74%	0.132	0.360
4	−0.000530	51.34%	0.011**	0.000***
5	−0.000042	50.18%	0.226	0.800
6	−0.000487	50.74%	0.026**	0.020**
7	−0.000562	51.49%	0.011**	0.000***
8	−0.000144	49.59%	0.579	0.440

can support H1b, indicating that sell signals have immediate predictive power and some short-term predictive ability post-signal release.

Figure 3 illustrates the distribution of SARs at event hour 0. For both buy and sell signals, SARs cluster sharply around zero, but with heavier tails and noticeable skew—positive for buy signals and negative for sell signals. Figure 4 shows cumulative abnormal returns (CAR), providing an aggregated view from -8 to $+8$ hours around the event. The CAR for buy signals steadily increases prior to the event and peaks at hour 0, suggesting price build-up in anticipation of signal issuance. After peaking, CAR flattens and slightly declines, reinforcing the interpretation of limited post-event momentum. For sell signals, the CAR curve exhibits a sharp decline at the event, then flattens with mild post-event decrease. These CAR dynamics confirm the asymmetric behavior observed in hourly AR measures: buy signals front-run small gains, while sell signals are more strongly associated with sharp, immediate losses.

We proceed to investigate whether trading signals have greater predictive power for specific types of cryptocurrencies, as outlined in hypotheses H2 through H4. This analysis focuses on event hour 0, where the main effect was previously established (see H1a and H1b). To assess whether the predictive strength of signals differs significantly across subsamples (e.g., low vs. high market capitalization), we employ a two-tailed t -test comparing abnormal returns (AR) at event hour 0 between groups. Table 8 presents the subsample analysis used to test H2 through H4, including statistical tests for differences in abnormal returns between groups. Appendix B reports the CARs over the event window for each subgroup to visualize the aggregated price dynamics.

Panel A of Table 8 examines the subsamples for market capitalization. For buy signals, the abnormal return (AR) is significantly higher for low-cap cryptocurrencies (0.006732) than for high-cap ones (0.003802), with a significant difference ($p < 0.01$). Both subgroups are strongly significant individually ($p < 0.01$), supporting H2a. For sell signals, a similar pattern emerges: low-cap coins exhibit more negative returns ($AR = -0.007722$) compared to high-cap coins ($AR = -0.005208$), with a significant difference

between them ($p < 0.01$). These results support H2b as well. Overall, the findings confirm H2a and H2b, suggesting that crowd-based trading signals are more predictive for cryptocurrencies with lower market capitalization.

Panel B investigates recent cryptocurrency performance. Buy signals yield higher abnormal returns for low-performance cryptocurrencies ($AR = 0.005778$) than high-performance ones ($AR = 0.004825$), with a significant difference ($p < 0.01$). Both subgroups are individually significant ($p < 0.01$), confirming H3a. A similar dynamic is observed for sell signals: AR is more negative for the low-performance group (-0.007066) than for the high-performance group (-0.005514), with the difference also significant ($p < 0.01$). These findings support H3b, indicating that signals tend to be more predictive for cryptocurrencies with recently weak performance. Taken together, the results confirm H3a and H3b, highlighting that crowd signals are particularly effective during periods of underperformance.

Panel C analyzes the impact of cryptocurrency age. Buy signals show nearly identical abnormal returns for older ($AR = 0.005918$) and younger ($AR = 0.005895$) cryptocurrencies, with no meaningful difference between the groups, despite both being highly significant individually ($p < 0.01$). This leads to a rejection of H4a. For sell signals, older cryptocurrencies exhibit slightly more negative AR ($AR = -0.007204$) than younger ones ($AR = -0.006923$), and the difference is marginally significant ($p < 0.1$), but in the opposite direction of the expected effect. Both subgroups remain individually significant ($p < 0.01$), but the direction of the difference leads us to reject H4b. Therefore, the results reject H4a and H4b, suggesting that cryptocurrency age is not strongly associated with the strength of abnormal returns around signals.

Exemplary trading strategy

Our event study demonstrates that social media signals, on average, exhibit predictive power, especially within the first hour after signal release. However, it remains an open question whether these signals can be utilized to construct a

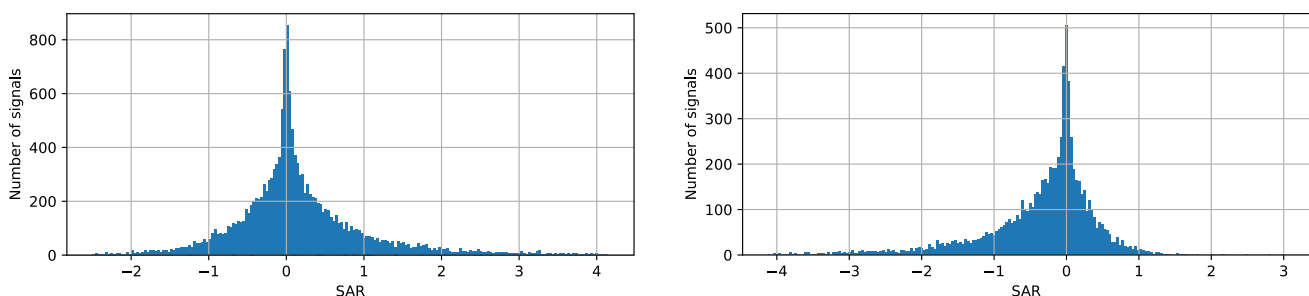


Fig. 3 Distribution of SAR for buy (left) and sell (right) signals in event hour 0

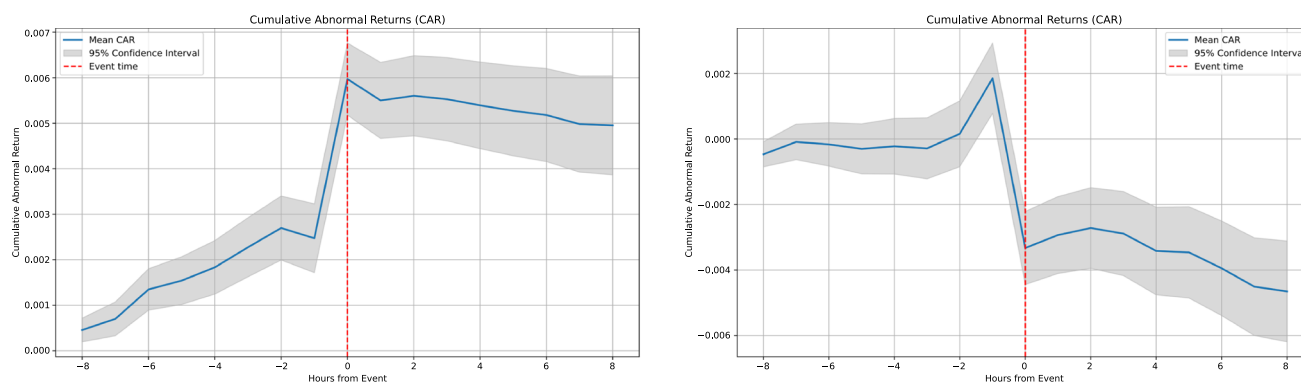


Fig. 4 Cumulative AR for buy (left) and sell (right) signals from -8 to $+8$ h around event

profitable trading strategy in practice. To explore this, we develop and simulate a simple hypothetical trading strategy and evaluate its performance on an out-of-sample dataset spanning July 1, 2023, to February 15, 2024. The signal extraction method is consistent with the “[Crowd-based trading signals](#)” section and yields a total of 747 buy signals. We restrict the strategy to buy-only trades, reflecting the practical reality that most cryptocurrency exchanges offer spot trading without the ability to short. Furthermore, we limit the tradable assets to cryptocurrencies in the intersection of the low market capitalization and low-performance categories. As shown in our subsample analysis in Table 8, these cryptocurrencies exhibit the most pronounced abnormal returns.

The trading strategy is straightforward: a position is opened at the asset’s hourly open price at the time of the signal and closed one hour later at the close price. We assume

fixed position sizing and apply a 0.3% transaction fee per trade. To mitigate the influence of extreme returns and potential data errors—especially common in illiquid or small-cap cryptocurrencies—we exclude the top and bottom 3% of trade returns based on a quantile filter. To aggregate performance, we compute the average return per day across all trades and construct an equal-weighted capital curve, avoiding compounding across individual trades. This approach assumes that capital is allocated evenly across all signals received each day, which reflects a practical execution framework under capital constraints.

Figure 5 plots the normalized capital curve of the strategy in comparison to two benchmarks: the S&P 500 index and the CCI30 crypto index, which tracks the 30 largest cryptocurrencies by market capitalization, excluding stablecoins. The inclusion of these benchmarks serves two purposes: the

Table 8 Subsample analysis. Significance of differences between subsamples (e.g., low vs high) based on two-sample t-tests. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Panel A: Cryptocurrency market capitalization

Subsample	Buy signals			Sell signals		
	Low	High	Diff. Sig	Low	High	Diff. Sig
AR	0.006732	0.003802	***	−0.007722	−0.005208	***
Z_{rank} p -value	0.000	0.000		0.000	0.000	
p_{perm} p -value	0.000	0.000		0.000	0.000	

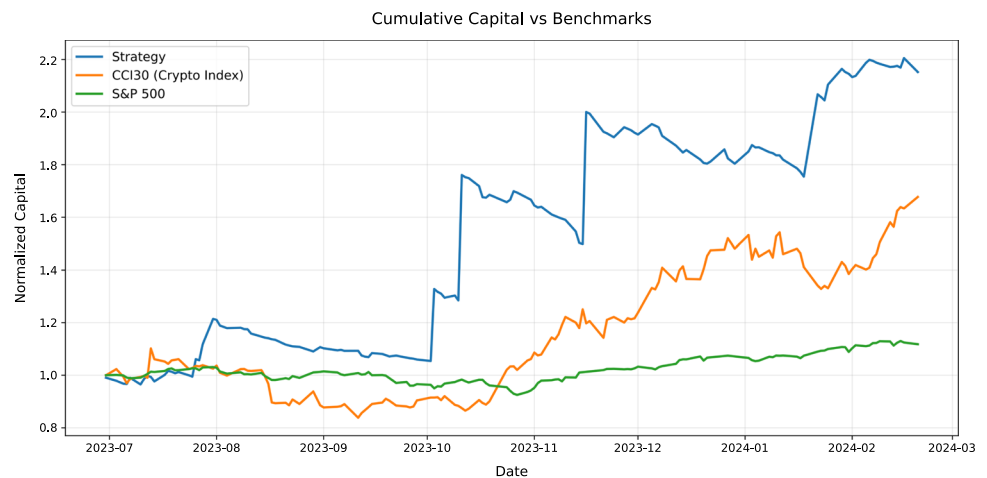
Panel B: Cryptocurrency performance

Subsample	Buy signals			Sell signals		
	Low	High	Diff. Sig	Low	High	Diff. Sig
AR	0.005778	0.004825	***	−0.007066	−0.005514	***
Z_{rank} p -value	0.000	0.000		0.000	0.000	
p_{perm} p -value	0.000	0.000		0.000	0.000	

Panel C: Cryptocurrency age

Subsample	Buy signals			Sell signals		
	Old	Young	Diff. Sig	Old	Young	Diff. Sig
AR	0.005918	0.005895		−0.007204	−0.006923	*
Z_{rank} p -value	0.000	0.000		0.000	0.000	
p_{perm} p -value	0.000	0.000		0.000	0.000	

Fig. 5 Cumulative returns of signal-based trading strategy and benchmarks on out-of-sample dataset



S&P 500 provides a traditional market reference point, while the CCI30 index reflects the broader performance of the crypto market over the same period. To evaluate the strategy's performance, we consider several key metrics, including total return, Sharpe ratio, maximum drawdown, and win rate (McNeil et al., 2015), summarized in Table 9. Overall, the strategy delivers strong returns with a solid risk-adjusted profile, outperforming both benchmarks in total return while maintaining a competitive Sharpe ratio. Despite a relatively low win rate, the strategy's performance is driven by a few highly profitable trades, indicating a positively skewed return distribution. Compared to the S&P 500 and the CCI30 index, the strategy demonstrates superior capital growth and controlled drawdowns.

It is important to emphasize that this strategy serves only as a baseline example of how the extracted signals could be utilized. In practice, one would likely incorporate more refined signal selection criteria, dynamically adjust allocations, and integrate additional data sources to improve robustness and execution.

Occurrence of signal bots

Bots are known to be highly active on financial social media, particularly on the platform X, where it is estimated that 1 to 14% of posts may originate from bot accounts (Kraaijeveld & De Smedt, 2020). Research has found that these bot-generated posts can influence financial markets, includ-

ing the stock market (Fan et al., 2020). Various types of bots have been studied in the literature, such as those engaging in cashtag piggybacking, where low-value assets are mentioned alongside high-value ones (Cresci et al., 2019) or mass-reposting bots (Tardelli et al., 2020). Given their presence in other areas of financial social media, we also expect bots in our signal dataset. During a qualitative review of the signals, we identified a distinct type of bot, which we term the *signal bot*. This bot frequently posts signals using a set of seemingly identical textual templates. For instance, typical posts include phrases like "\$SOL 1 minute - Entry Signal Time: 8/8 13:8..." or "\$GALA 1 minute - Entry Signal Time: 27/7 18:1..." Such signals appear to be automatically generated based on some automated trading strategies that trigger posting of signals.

To detect such occurrences, we developed a heuristic approach based on textual content analysis. Initially, we performed pairwise comparisons of posts from each user to identify templates. This was accomplished using the `SequenceMatcher` class from Python's `difflib` module, which calculates a similarity ratio based on matching sequences of text. We set a similarity threshold of 0.7, whereby posts with a similarity ratio above this threshold were considered to have been generated from the same template. For each user, if a significant proportion of their text pairs (exceeding 0.5) surpassed this similarity threshold, the user was classified as a bot. In addition to the signal bot detection algorithm, we applied a heuristic-based method adapted from prior research (Kraaijeveld & De Smedt, 2020) that classifies a post as potentially bot-generated if it meets multiple criteria, such as containing phrases like "give away" or "giving away," mentioning "pump" alongside "register" or "join" (often linked to fraudulent pump-and-dump schemes), including more than 5 hashtags, featuring more than 5 cryptocurrency ticker or having a platform user name containing the word "bot."

Table 9 Key performance metrics of trading strategy and benchmarks

Metric	Strategy	CCI30	S&P 500
Total return	115.23%	67.72%	11.80%
Sharpe ratio	2.43	2.60	2.02
Max drawdown	-14.92%	-23.91%	-10.28%
Win rate	27.97%	57.34%	56.64%

Our analysis, conducted on the unbalanced dataset of signals, suggests that around 6% of the accounts can be classified as signal bots. These accounts likely use some trading strategy and post automatically. When using the heuristic by Kraaijeveld and De Smedt (2020), 5% of accounts are identified as bots. We therefore conclude that bots are active in the signal domain, but their presence is not unusually high relative to overall financial social media. We consider these signals as part of the broader phenomenon and conclude that they have not significantly impacted our study.

Robustness checks

As part of our zero-shot classification approach, we also identified a category of neutral signals—posts that did not advise an explicit buy or sell recommendation. To assess the robustness of our findings, we conducted the same event study analysis using these neutral signals. This provides a robustness check to assess whether the return patterns linked to buy and sell signals are driven by their directional content, or if similar effects also occur around non-direction signal posts. The full results are provided in Appendix C. The CAR around neutral signals remains flat across the event window, with no clear movement at the event hour. The analysis of abnormal returns in the individual hours further confirms this: the abnormal return at hour 0 is small and statistically insignificant ($AR = -0.000324$), and no meaningful or consistent effects are found in surrounding hours. These results support the interpretation that the observed effects are associated with the directional nature of the signals, rather than reflecting general market responses to any signal-related communication.

To test the robustness of our event study, we re-estimated abnormal returns using a market-adjusted return model, which subtracts the return of the capitalization-weighted crypto index from the contemporaneous return of each asset, avoiding the assumptions of regression-based models. Our main results remained consistent with those obtained under CAPM, confirming that our findings are not model-dependent. While CAPM was developed for equity markets and may not fully reflect the volatility and dynamics, or absence of a risk-free rate in crypto markets, it offers an advantage by adjusting for each asset's historical sensitivity to the market through beta. The consistency of results across both models supports the robustness of our main conclusions.

Discussion

This study explores the predictive power of crowd-based social media signals in the context of cryptocurrencies. Our central finding is that directional social media signals are associated with short-term abnormal returns in the

direction of the signal. Compared to prior sentiment-based approaches (Subramanian et al., 2023), our method provides an alternative by directly leveraging the explicit intent of crowd-generated trading signals rather than relying on aggregated sentiment scores. However, we view this approach not as replacement, but as a potential supplement to existing sentiment-based or other trading strategies, particularly in high-frequency contexts where actionable intent may provide more immediate market signals. The strongest effects are concentrated at the moment of signal release (event hour 0). This contrasts with social media signals in stock markets, where predictive power often unfolds over longer time horizons such as days or months (Buz & De Melo, 2024). Additionally, some signals appear to follow, rather than lead, market movements—suggesting partially reactive behavior by users, consistent with findings from the Bitcoin market (Xie, 2022). We also observe that buy signals are followed by slightly negative returns a few hours after the event, which may reflect short-term overreactions or potential manipulation dynamics such as pump-and-dump activity.

The predictive power of these signals also varies with cryptocurrency characteristics. We observe that buy and sell signals are more predictive for low market capitalization cryptocurrencies than for larger ones. This suggests that smaller assets, which may be less efficient or more prone to behavioral trading, are more strongly associated with crowd signals. However, even large-cap cryptocurrencies show statistically significant effects, indicating that directional crowd signals are predictive across the broader market.

Performance level is also associated with differences in predictive power. Signals associated with lower-performing cryptocurrencies show stronger effects than those linked to higher-performing ones. This indicates that crowd signals may be particularly useful during periods of weak market performance when informational inefficiencies are more likely to emerge. Both buy and sell signals show this pattern, reinforcing the idea that underperforming cryptocurrencies tend to show greater responsiveness to crowd signals.

Finally, we find little evidence that cryptocurrency age is strongly associated with differences in signal performance. Buy signals yield nearly identical abnormal returns for both older and younger cryptocurrencies, while sell signals show a slightly stronger effect for older assets, though in the opposite direction of the initial hypothesis. Overall, the results indicate that the behavior of crowd signals is broadly consistent across both established and emerging cryptocurrencies.

The results of the hypothetical trading strategy, based on an out-of-sample dataset, demonstrate that these signals can generate profitable returns in real-world trading scenarios. However, it is essential to emphasize that these findings should not be interpreted as specific trading advice. The primary goal of the strategy was to illustrate the potential magnitude of returns that could be achieved. Factors such

as slippage were not accounted for in our analysis. There are several ways to further refine and enhance this strategy. For instance, improving the signal filtering mechanisms can enhance practical applications. In practice, it is unlikely that such signals would be used in isolation; rather, they can serve as a valuable component within a more sophisticated and comprehensive trading system.

We propose a potential explanation for our findings that begins with significant market events, such as a rapid increase or decrease in the value of a cryptocurrency. These movements could be triggered by various factors unrelated to social media, such as influential news (or even fake news (Clarke et al., 2021)), strategic moves by large traders, or actions by algorithmic traders. Following these market events, social media signals are generated and disseminated by real users and bots. These signals appear to be reactive in nature, suggesting that they are likely based on the observation of factors unrelated to social media. Following the release of these signals, we typically see a continuation of the initial market trend, whether an upward rally or a downward sell-off, indicating a short-term autocorrelation of returns. However, this trend often reverses afterward. This pattern suggests that the market's initial reaction preceding social media signals is ephemeral. While these signals may not provide deep market insight or inherent predictive power, they do reflect short-term market behavior. The immediate trend continuation following a signal release presents a temporary trading opportunity, prior to the market's self-correction. The positive abnormal return for buy signals indicates a slight tendency to outperform following these signals, whereas the negative abnormal return for sell signals suggests a slight underperformance.

Theoretical contributions

This study offers three key contributions that extend and refine existing theoretical perspectives in these domains. First, we re-conceptualize the use of social media data in algorithmic trading by moving beyond traditional sentiment analysis toward the extraction of explicit crowd-sourced trading intent. Prior literature on sentiment-based trading has primarily focused on aggregating sentiment from textual data, often treating sentiment as a proxy for investor mood or market direction (Bollen et al., 2011; Subramanian et al., 2023). However, sentiment is inherently noisy and frequently detached from concrete behavioral intent. Our findings qualify the wisdom of crowds perspective (Surowiecki, 2005) by demonstrating that actionable collective intelligence in volatile cryptocurrency markets can be efficiently derived from disaggregated, intent-based signals. This extends the application of crowd wisdom to environments where not all crowd contributions are equally valuable, demonstrating that effective use of crowd intelligence does not require sentiment

aggregation but can instead rely on direct analysis of individual signals. Moreover, by focusing on explicit trading signals rather than broad sentiment aggregation, our study suggests that collective intelligence in these noisy environments might be more effectively captured by filtering for specific, high-intent contributions rather than simply averaging all available crowd inputs.

Second, we contribute to the understanding of social media as a distributed and dynamic financial information infrastructure, particularly within emerging digital asset markets characterized by high information asymmetry (Alfieri et al., 2025). Unlike traditional stock markets, cryptocurrency ecosystems lack centralized intermediaries and institutional analysts (Joebges et al., 2025), making them particularly susceptible to decentralized forms of knowledge production and dissemination. Our findings provide empirical support for the notion that general-purpose social media platforms such as X, Reddit, Stocktwits, and Telegram function as systems, where market-relevant information emerges from collective discourse. This phenomenon could be understood through an extension of signaling theory. The explicit signals we identify serve as observable indicators attempting to reduce the information asymmetries between actors. The predictive power of these signals confirms their effectiveness.

Third, our identification of a distinct class of signal bots raises important theoretical questions about algorithmic agency and digital trust. While much of the prior work on bots in financial contexts has emphasized manipulation or misinformation (Tardelli et al., 2022), we observe bots that provide content that can be perceived as useful. This challenges existing classifications of financial bot activity as harmful or deceptive and suggests that algorithmic entities can play an active, even constructive, role in decentralized financial ecosystems and contribute to the crowd wisdom.

Practical implications

Our research yields several practical implications for different stakeholders operating within the cryptocurrency and financial data ecosystem. First, for institutional investors, our findings support the integration of social media-based trading signals as a legitimate alternative data source. We demonstrate that, even in an unstructured and noisy environment, meaningful short-term predictive insights can be extracted using LLMs in a zero-shot setting with an alternative approach to sentiment analysis. This approach can enable investors to efficiently incorporate real-time crowd intelligence into their trading algorithms, risk models, or market surveillance systems—particularly relevant for stakeholders such as hedge funds already engaging in high-frequency strategies based on social media data (Deng et al., 2018). For retail investors, however, the practical utility of such signals is more limited. Given the narrow time window in which

effects materialize, these signals may be difficult to act on in time and could even mislead, especially in highly volatile markets.

Second, for social media platforms that serve as hubs for financial discussion, our study highlights an emerging content category: public trading signals. Platforms may benefit from recognizing and structuring this content differently, for example, by developing tagging systems, filters, or dedicated spaces for signal sharing. These enhancements may also increase the usefulness of such signals for retail investors and foster more transparent discourse. However, as financial content becomes more prominent, platforms also carry an increasing responsibility to ensure that such activity complies with legal and regulatory boundaries, especially when trading advice is disseminated to large audiences. The need to distinguish between personal opinions, coordinated activity, and potentially manipulative behavior places pressure on platforms to develop both technical tools and policy mechanisms in line with regulatory standards. Unlike dedicated platforms such as eToro (Pan et al., 2012), where social trading is explicitly structured, trading advice on general-purpose platforms is less curated and often escapes formal oversight, making governance and accountability significantly more challenging.

Third, for regulators and policymakers, the widespread presence of public trading signals on social media underscores the need for clearer guidance and frameworks suited to this evolving informational environment. This can include requirements for disclosure, bot identification, and content verification. Regulators are in a position to define and enforce standards suited to this new landscape, especially in the less regulated cryptocurrency markets. To this end, traditional stock market surveillance tools (Sio, 2024) must be adapted to capture the specific dynamics of cryptocurrencies, including the large presence of direct trading advice.

Limitations and future work

This study has some limitations that present opportunities for future research. One limitation is the source bias in our data collection. The majority of our signals were derived from X, reflecting its prominence in cryptocurrency discussions. This reliance on a single source could introduce bias, as the predictive power of signals from other social media platforms might offer different insights, as suggested by prior research comparing cross-platform sentiment effects (Mai et al., 2018). As a result, our findings should be interpreted with the understanding that they, to a large extent, reflect the dynamics of the platform X. Furthermore, our study was constrained by data limitations, particularly the hourly resolution of price data. This granularity may not fully capture the rapid fluctuations characteristic of the cryptocurrency market and social media signals. A finer temporal resolu-

tion (e.g., minutely) could potentially reveal more nuanced insights into the predictive power of these signals and should be considered for practical applications (similar observations have been made, for example, by Deng et al. (2018)). A further limitation relates to the use of CAPM in cryptocurrency markets, where some assumptions of the model may not fully hold. As a result, absolute abnormal returns should be interpreted cautiously, as they may partly reflect limitations of the model. Finally, we applied a signal threshold to balance the dataset by retaining only cryptocurrencies with at least 100 signals and subsampling to a fixed number per asset. While this improves comparability between cryptocurrencies, it introduces a non-random sampling process that distorts the underlying signal distribution. As a result, this limits generalizability, particularly when applying the findings to real-world settings where signal volume is heavily skewed toward a few dominant assets.

Looking ahead, an important direction for future research is the integration of trading signals into full trading systems (Subramanian et al., 2023). While we focused on establishing the predictive potential of signals, operationalizing them within real-time, high-frequency trading frameworks remains an open challenge. This would require not only methodological refinement in signal extraction but also adaptive mechanisms to evaluate and weigh signals dynamically. In that regard, the imbalance of signals for various cryptocurrencies needs to be accounted for. Future studies could explore a more differentiated treatment of signals by considering the identity and influence of the source. For instance, signals from high-following or verified accounts may differ in impact from those posted by lesser-known users or signal bots. Furthermore, our identification of signal bots points to a particularly promising future research direction in understanding their various schemes and incentives. Theoretically, future work could also examine herding effects, which have been studied in the context of financial markets and cryptocurrencies (Youssef, 2022; King & Koutmos, 2021) and may also be at play in the signal domain, where signals could depend on each other through mechanisms such as information cascades. Finally, while our findings show stronger associations between signals and returns for low market capitalization or poorly performing assets, this may in part reflect underlying risk premia rather than informational inefficiency. Future work should aim to disentangle predictive value from compensation for elevated risk in these market segments.

Conclusion

This study explores cryptocurrency trading through the use of crowd-based social media signals. We show that these signals, which reflect the collective intelligence of social media communities, have significant predictive power on

the cryptocurrency markets. The results indicate that social media signals are not just noise, but contain valuable insights that can aid decision-making for cryptocurrency trading systems. Through the examination of various cryptocurrencies and the incorporation of signals from multiple social media platforms, we confirm the theory of the wisdom of the crowds in the context of cryptocurrency trading. Our analysis emphasizes the immediate predictive power of these signals, particularly at the time of their release, highlighting their usefulness for high-frequency trading decisions. However, our research also acknowledges the nuances and complexity of the cryptocurrency market. Although the results are robust across various cryptocurrencies, variations in predictive power are observed based on factors such as market capitalization and performance.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12525-025-00815-6>.

Acknowledgements We thank Stockpulse (<https://stockpulse.ai/>) for providing the essential research data for this study.

Author contribution Frederic Haase: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, project administration.

Tom Celig: methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review and editing, visualization.

Oliver Rath: conceptualization, investigation, writing—original draft, writing—review and editing, visualization, project administration.

Detlef Schoder: supervision, project administration, funding acquisition, writing—review editing.

Funding Open Access funding enabled and organized by Projekt DEAL. This work is part of the research project AFFIN funded by the German Federal Ministry of Education and Research (Grant no.: 01IS21045B).

Data Availability Our sample of signals is available upon request.

Declarations

Ethical approval Not applicable

Consent to participate Not applicable

Consent for publication Not applicable

Conflict of interest The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process During the preparation of this work, the authors used *Grammarly* in order to improve language and readability. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adap-

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

References

- Agrawal, P., Buz, T., & Melo, G. D., (2022). WallStreetBets Beyond GameStop, YOLOs, and the Moon: The unique traits of Reddit's finance communities. *AMCIS 2022 Proceedings* https://aisel.aisnet.org/amcis2022/sig_sc/sig_sc/8
- Alfieri, E., Burlacu, R., & Enjolras, G., (2025). Cryptocurrency bubbles, information asymmetry and noise trading. *Journal of Risk Finance*, 26, 295–319. <https://ideas.repec.org/a/eme/jrfpps/jrf-07-2024-0220.html>
- Aloosh, A., Ouzan, S., & Shahzad, S. J. H., (2022). Bubbles across meme stocks and cryptocurrencies. *Finance Research Letters*, 49, 103155. <https://linkinghub.elsevier.com/retrieve/pii/S1544612322003774>, <https://doi.org/10.1016/j.frl.2022.103155>
- Alt, R., Fridgen, G., & Chang, Y., (2024). The future of fintech-Towards ubiquitous financial services. *Electronic Markets*, 34, 3, s12525-023-00687-8. <https://doi.org/10.1007/s12525-023-00687-8>
- Anh, N.P., & Wolf, M., (2023). A note on testing AR and CAR for event studies. Working Paper. Working Paper. <https://www.econstor.eu/handle/10419/268859>, <https://doi.org/10.5167/uzh-229116>
- Ante, L., (2023). How Elon Musk's Twitter activity moves cryptocurrency markets. *Technological Forecasting and Social Change*, 186, 122112. <https://www.sciencedirect.com/science/article/pii/S0040162522006333>, <https://doi.org/10.1016/j.techfore.2022.122112>
- Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance*, 59, 1259–1294. <https://doi.org/10.1111/j.1540-6261.2004.00662.x>
- Apesteguia, J., Oechssler, J., & Weidenholzer, S. (2020). Copy trading. *Management Science*, 66, 5608–5622. <https://doi.org/10.1287/mnsc.2019.3508>
- Baker, M., & Wurgler, J., (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21, 129–152. <https://www.aeaweb.org/articles?id=10.1257/jep.21.2.129>, <https://doi.org/10.1257/jep.21.2.129>
- Bari, O. A., & Agah, A., (2020). Ensembles of text and time-series models for automatic generation of financial trading signals from social media content. *Journal of Intelligent Systems*, 29, 753–772. <https://www.degruyter.com/document/doi/10.1515/jisys-2017-0567/html>, <https://doi.org/10.1515/jisys-2017-0567>
- Beck, R., Müller-Bloch, C., & King, J., (2018). Governance in the blockchain economy: A framework and research agenda. *Journal of the Association for Information Systems*, 19. <https://aisel.aisnet.org/jais/vol19/iss10/1>
- Beinke, M., Beinke, J. H., Anton, E., & Teuteberg, F. (2024). Breaking the chains of traditional finance: A taxonomy of decentralized finance business models. *Electronic Markets*, 34, 29. <https://doi.org/10.1007/s12525-024-00704-4>
- Bollen, J., Mao, H., & Zeng, X., (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1–8. <https://www.sciencedirect.com/science/article/>

- pii/S187775031100007X, <https://doi.org/10.1016/j.jocs.2010.12.007>
- Bose, I., & Pal, R., (2012). Do green supply chain management initiatives impact stock prices of firms? *Decision Support Systems*, 52, 624–634. <https://www.sciencedirect.com/science/article/pii/S0167923611002004>, <https://doi.org/10.1016/j.dss.2011.10.020>
- Bouri, E., Kamal, E., & Kinader, H., (2023). FTX Collapse and systemic risk spillovers from FTX Token to major cryptocurrencies. *Finance Research Letters*, 56, 104099. <https://linkinghub.elsevier.com/retrieve/pii/S1544612323004713>, <https://doi.org/10.1016/j.frl.2023.104099>
- Breitmayer, B., Massari, F., & Pelster, M., (2019). Swarm intelligence? Stock opinions of the crowd and stock returns. *International Review of Economics & Finance*, 64, 443–464. <https://linkinghub.elsevier.com/retrieve/pii/S1059056018305057>, <https://doi.org/10.1016/j.iref.2019.08.006>
- Buz, T., & De Melo, G. (2024). Democratisation of retail trading: A data-driven comparison of Reddit's WallStreetBets to investment bank analysts. *Journal of Business Analytics*, 1–17. <https://doi.org/10.1080/2573234X.2024.2354191>
- Celig, T., Milosevic, L., Ockenga, T.A., & Schoder, D., (2024). Linking social and wealth metrics to portfolio outcomes in DeFi: A statistical analysis of on-chain portfolio data. *ICIS 2024 Proceedings*, <https://aisel.aisnet.org/icis2024/blockchain/blockchain/4>
- Chen, H., Yu Jeffrey, H., & Huang, S. (2019). Monetary incentive and stock opinions on social media. *Journal of Management Information Systems*, 36, 391–417. <https://doi.org/10.1080/07421222.2019.1598686>
- Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27, 1367–1403. <https://doi.org/10.1093/rfs/hhu001>
- Chung, D., & Hrazdil, K. (2010). Liquidity and market efficiency: Analysis of NASDAQ firms. *Global Finance Journal*, 21, 262–274. <https://doi.org/10.1016/j.gfj.2010.09.004>
- Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2021). Fake news, investor attention, and market reaction. *Information Systems Research*, 32, 35–52. <https://doi.org/10.1287/isre.2019.0910>
- Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37, 39–67. <https://doi.org/10.1177/0149206310388419>
- Cresci, S., Lillo, F., Regoli, D., Tardelli, S., & Tesconi, M. (2019). CASHtag piggybacking: Uncovering spam and bot activity in stock microblogs on Twitter. *ACM Trans. Web.*, 13, 11:1–11:27. <https://doi.org/10.1145/3313184>
- Deng, S., Huang, Z. J., Sinha, A. P., & Zhao, H., (2018). The interaction between microblog sentiment and stock returns: An empirical examination. *MIS Quarterly*, 42, 895–A13. <https://www.jstor.org/stable/26635058>
- Dewally, M. (2003). Internet investment advice: Investing with a rock of salt. *Financial Analysts Journal*, 59, 65–77. <https://doi.org/10.2469/faj.v59.n4.2546>
- Dong, L., Zheng, H., Li, L., & Zhou, C., (2024). Distilling wisdom of crowds in online communities: A novel prediction market constructed with comment posters. *Decision Support Systems*, 180, 114190. <https://linkinghub.elsevier.com/retrieve/pii/S016792362400023X>, <https://doi.org/10.1016/j.dss.2024.114190>
- Dos Santos, B. L., Peffers, K., & Mauer, D. C. (1993). The impact of information technology investment announcements on the market value of the firm. *Information Systems Research*, 4, 1–23. <https://doi.org/10.1287/isre.4.1.1>
- Egan, J., (2023). Nearly 80% Of young adults get financial advice from this surprising place. <https://www.nasdaq.com/articles/nearly-80-of-young-adults-get-financial-advice-from-this-surprising-place>. Accessed 16 June 2025.
- Fama, E. F., (1965). The behavior of stock-market prices. *The Journal of Business*, 38, 34. <https://www.jstor.org/stable/2350752>, <https://doi.org/10.1086/294743>
- Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 18, 25–46. <https://doi.org/10.1257/0895330042162430>
- Fan, R., Talavera, O., & Tran, V. (2020). Social media bots and stock markets. *European Financial Management*, 26, 753–777. <https://doi.org/10.1111/eufm.12245>
- Garcia, D., & Schweitzer, F. (2015). Social signals and algorithmic trading of Bitcoin. <https://doi.org/10.1098/rsos.150288>
- García-Corral, F. J., Cordero-García, J. A., de Pablo-Valenciano, J., & Uribe-Toril, J. (2022). A bibliometric review of cryptocurrencies: How have they grown? *Financial Innovation*, 8, 2. <https://doi.org/10.1186/s40854-021-00306-5>
- Gary, G., (2024). SEC: Statement on the Approval of Spot Bitcoin Exchange-Traded Products. <https://www.sec.gov/news/statement/gensler-statement-spot-bitcoin-011023>. Accessed 16 June 2025.
- Gerritsen, D. F., Lugtigheid, R. A., & Walther, T., (2022). Can Bitcoin investors profit from predictions by crypto experts? *Finance Research Letters*, 46, 102266. <https://linkinghub.elsevier.com/retrieve/pii/S1544612321003081>, <https://doi.org/10.1016/j.frl.2021.102266>
- Gjerstad, P., Meyn, P. F., Molnár, P., & Næss, T. D., (2021). Do President Trump's tweets affect financial markets? *Decision Support Systems*, 147, 113577. <https://www.sciencedirect.com/science/article/pii/S0167923621000877>, <https://doi.org/10.1016/j.dss.2021.113577>
- Gottschlich, J., & Hinz, O., (2014). A decision support system for stock investment recommendations using collective wisdom. *Decision Support Systems*, 59, 52–62. <https://www.sciencedirect.com/science/article/pii/S0167923613002522>, <https://doi.org/10.1016/j.dss.2013.10.005>
- Gwet, K. L., (2014). *Handbook of Inter-Rater Reliability, 4th Edition: The Definitive Guide to Measuring The Extent of Agreement Among Raters*. Advanced Analytics, LLC. Google-Books-ID: fac9BQAAQBAJ.
- Hamrick, J., Rouhi, F., Mukherjee, A., Vasek, M., Moore, T., & Gandal, N., (2021). Analyzing Target-Based Cryptocurrency Pump and Dump Schemes. In: *Proceedings of the 2021 ACM CCS Workshop on Decentralized Finance and Security*. Association for Computing Machinery, New York, NY, USA. pp. 21–27. <https://doi.org/10.1145/3464967.3488591>
- Haase, F., Rath, O., Krauß, J., & Schoder, D. (2025). The role of influencers in shaping crowd sentiment. *Business & Information Systems Engineering*. <https://doi.org/10.1007/s12599-025-00947-1>
- Henderson, G. V., (1990). Problems and Solutions in Conducting Event Studies. *The Journal of Risk and Insurance*, 57, 282–306. <https://www.jstor.org/stable/253304>, <https://doi.org/10.2307/253304>
- Hägele, S., (2024). Centralized exchanges vs. decentralized exchanges in cryptocurrency markets: A systematic literature review. *Electronic Markets*, 34, 33. <https://doi.org/10.1007/s12525-024-00714-2>
- Im, K. S., Dow, K. E., & Grover, V. (2001). Research Report: A Reexamination of IT Investment and the Market Value of the Firm-An Event Study Methodology. *Information Systems Research*, 12, 103–117. <https://doi.org/10.1287/isre.12.1.103.9718>
- Jalal, R. N. U. D., Alon, I., & Paltrinieri, A. (2021). A bibliometric review of cryptocurrencies as a financial asset. *Technology Analysis & Strategic Management*, 1–16. <https://doi.org/10.1080/09537325.2021.1939001>
- Joebges, H., Herr, H., & Kellermann, C. (2025). Crypto assets as a threat to financial market stability. *Eurasian Economic Review*. <https://doi.org/10.1007/s40822-025-00311-4>

- Kamps, J., & Kleinberg, B. (2018). To the moon: defining and detecting cryptocurrency pump-and-dumps. *Crime Science*, 7, 18. <https://doi.org/10.1186/s40163-018-0093-5>
- Kang, H. J., Lee, S. G., & Park, S. Y. (2022). Information Efficiency in the Cryptocurrency Market: The Efficient-Market Hypothesis. *Journal of Computer Information Systems*, 62, 622–631. <https://doi.org/10.1080/08874417.2021.1872046>
- Kaplanski, G., & Levy, H. (2010). Sentiment and stock prices: The case of aviation disasters. *Journal of Financial Economics*, 95, 174–201. <https://www.sciencedirect.com/science/article/pii/S0304405X09002086>, <https://doi.org/10.1016/j.jfineco.2009.10.002>
- Kim, K., Lee, S. Y. T., & Kauffman, R. J. (2023). Social informedness and investor sentiment in the GameStop short squeeze. *Electronic Markets*, 33, 23. <https://doi.org/10.1007/s12525-023-00632-9>
- King, T., & Koutmos, D. (2021). Herding and feedback trading in cryptocurrency markets. *Annals of Operations Research*, 300, 79–96. <https://doi.org/10.1007/s10479-020-03874-4>
- Kolari, J. W., & Pynnonen, S. (2011). Nonparametric rank tests for event studies. *Journal of Empirical Finance*, 18, 953–971. <https://www.sciencedirect.com/science/article/pii/S0927539811000624>, <https://doi.org/10.1016/j.jempfin.2011.08.003>
- Kraaijeveld, O., & De Smedt, J. (2020). The predictive power of public Twitter sentiment for forecasting cryptocurrency prices. *Journal of International Financial Markets, Institutions and Money*, 65. <https://ideas.repec.org/a/eee/intfin/v65y2020ics104244312030072x.html>
- La Morgia, M., Mei, A., Sassi, F., Stefa, J., 2023. The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations. *ACM Transactions on Internet Technology*, 23, 11:1–11:28. <https://doi.org/10.1145/3561300>
- Lachana, I., & Schröder, D. (2025). Investor sentiment and stock returns: Wisdom of crowds or power of words? Evidence from Seeking Alpha and Wall Street Journal. *Journal of Financial Markets*, 74, 100970. <https://www.sciencedirect.com/science/article/pii/S1386418125000102>, <https://doi.org/10.1016/j.finmar.2025.100970>
- Larrick, R. P., Mannes, A. E., & Soll, J. B. (2012). The social psychology of the wisdom of crowds. In: *Social judgment and decision making*. Psychology Press, pp. 227–242.
- Levich, R., Conlon, T., & Potí, V., (2019). Measuring excess-predictability of asset returns and market efficiency over time. *Economics Letters*, 175, 92–96. <https://www.sciencedirect.com/science/article/pii/S0165176518305147>, <https://doi.org/10.1016/j.econlet.2018.12.022>
- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., Zettlemoyer, L., 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension, in: Jurafsky, D., Chai, J., Schluter, N., Tetreault, J. (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online. pp. 7871–7880. <https://aclanthology.org/2020.acl-main.703/>, <https://doi.org/10.18653/v1/2020.acl-main.703>
- Li, T., van Dalen, J., & van Rees, P. J. (2018). More than just Noise? Examining the Information Content of Stock Microblogs on Financial Markets. *Journal of Information Technology*, 33, 50–69. <https://doi.org/10.1057/s41265-016-0034-2>
- Li, X., & Whinston, A. B. (2020). Analyzing Cryptocurrencies. *Information Systems Frontiers*, 22, 17–22. <https://doi.org/10.1007/s10796-019-09966-2>
- Mai, F., Shan, Z., Bai, Q., Wang, X. S., & Chiang, R. H. (2018). How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis. *Journal of Management Information Systems*, 35, 19–52. <https://doi.org/10.1080/07421222.2018.1440774>
- McNeil, A.J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: concepts, techniques and tools*. Princeton series in finance. revised edition ed., Princeton University Press, Princeton, NJ. OCLC: ocn894625411.
- Meyer, E. A., Welp, I. M., & Sandner, P. (2024). Testing the credibility of crypto influencers: An event study on Bitcoin. *Finance Research Letters*, 60, 104864. <https://linkinghub.elsevier.com/retrieve/pii/S1544612323012369>, <https://doi.org/10.1016/j.frl.2023.104864>
- Momtaz, P. P. (2021). The Pricing and Performance of Cryptocurrency. *The European Journal of Finance*, 27, 367–380. <https://doi.org/10.1080/1351847X.2019.1647259>
- Nghiem, H., Muric, G., Morstatter, F., & Ferrara, E., (2021). Detecting cryptocurrency pump-and-dump frauds using market and social signals. *Expert Systems with Applications*, 182, 115284. <https://linkinghub.elsevier.com/retrieve/pii/S0957417421007156>, <https://doi.org/10.1016/j.eswa.2021.115284>
- Nofer, M., & Hinz, O. (2014). Are crowds on the internet wiser than experts? The case of a stock prediction community. *The Journal of Business Economics*, 84, 303–338. <https://www.proquest.com/docview/1932046447/abstract/5E49A24BB1AA4B3EPQ/1>, <https://doi.org/10.1007/s11573-014-0720-x>. num Pages: 303-338 Place: Heidelberg, Netherlands Publisher: Springer Nature B.V.
- Nofer, M., & Hinz, O. (2015). Using Twitter to Predict the Stock Market. *Business & Information Systems Engineering*, 57, 229–242. <https://doi.org/10.1007/s12599-015-0390-4>
- Pan, W., Altshuler, Y., & Pentland, A. (2012). Decoding Social Influence and the Wisdom of the Crowd in Financial Trading Network. In: *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing*, pp. 203–209. <https://ieeexplore.ieee.org/abstract/document/6406285>, <https://doi.org/10.1109/SocialCom-PASSAT.2012.133>
- Peterson, R. L., (2016). *Trading on Sentiment: The Power of Minds Over Markets*. John Wiley & Sons. Google-Books-ID: IOhCgAAQBAJ.
- Phillip, A., Chan, J. S. K., & Peiris, S. (2018). A new look at Cryptocurrencies. *Economics Letters*, 163, 6–9. <https://www.sciencedirect.com/science/article/pii/S0165176517304731>, <https://doi.org/10.1016/j.econlet.2017.11.020>
- Piñero-Chousa, J., Šević, A., & González-López, I. (2023). Impact of social metrics in decentralized finance. *Journal of Business Research*, 158, 113673. <https://www.sciencedirect.com/science/article/pii/S0148296323000310>, <https://doi.org/10.1016/j.jbusres.2023.113673>
- Pocher, N., Zichichi, M., Merizzi, F., Shafiq, M. Z., & Ferretti, S. (2023). Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics. *Electronic Markets*, 33, 37. <https://doi.org/10.1007/s12525-023-00654-3>
- pwc & AIMA. (2024). *6th Annual Global Crypto Hedge Fund Report. Technical Report 6th Annual Global Crypto Hedge Fund Report*. <https://www.pwc.com/gx/en/industries/financial-services/crypto-services/sixth-annual-global-crypto-hedge-fund-report.html>. Accessed 16 June 2025.
- Schwiderowski, J., Pedersen, A. B., & Beck, R. (2024). Crypto Tokens and Token Systems. *Information Systems Frontiers*, 26, 319–332. <https://doi.org/10.1007/s10796-023-10382-w>
- See-To, E. W. K., & Yang, Y. (2017). Market sentiment dispersion and its effects on stock return and volatility. *Electronic Markets*, 27(3), 283–296. <https://doi.org/10.1007/s12525-017-0254-5>
- Shanaev, S., Shuraeva, A., Vasenin, M., & Kuznetsov, M. (2019). Cryptocurrency Value and 51% Attacks: Evidence from Event Studies. *The Journal of Alternative Investments*, 22, 65–77. <https://doi.org/10.3905/jai.2019.1.081>
- Sigaki, H. Y. D., Perc, M., & Ribeiro, H. V. (2019). Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market. *Scientific Reports*, 9, 1440. <https://www.nature.com/>

- articles/s41598-018-37773-3, <https://doi.org/10.1038/s41598-018-37773-3>
- Sio, T., 2024. *Social Media Monitoring with Market Surveillance Tools*. <https://www.nasdaq.com/articles/social-media-monitoring-with-market-surveillance-tools>. Accessed 16 June 2025.
- Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87, 355. <https://doi.org/10.2307/1882010>
- Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *American Economic Review*, 92, 434–459. <https://doi.org/10.1257/00028280260136200>
- Stotz, O., Wanzenried, G., & Döhnert, K. (2010). Do fundamental indexes produce higher risk-adjusted returns than market cap indexes? Evidence for European stock markets. *Financial Markets and Portfolio Management*, 24, 219–243. <https://doi.org/10.1007/s11408-010-0135-9>
- Subramanian, H., Angle, P., Rouxelin, F., & Zhang, Z. (2023). A decision support system using signals from social media and news to predict cryptocurrency prices. *Decision Support Systems*, 114129. <https://linkinghub.elsevier.com/retrieve/pii/S016792362300204X>, <https://doi.org/10.1016/j.dss.2023.114129>
- Surowiecki, J., (2005). *The Wisdom of Crowds*. Knopf Doubleday Publishing Group. Google-Books-ID: hHUsHOHqVzEC.
- Talmor, E. (1981). Asymmetric Information, Signaling, and Optimal Corporate Financial Decisions. *The Journal of Financial and Quantitative Analysis*, 16, 413. <https://www.jstor.org/stable/2330363?origin=crossref>, <https://doi.org/10.2307/2330363>
- Tardelli, S., Avvenuti, M., Tesconi, M., Cresci, S., 2020. *Characterizing Social Bots Spreading Financial Disinformation*, Springer International Publishing, Cham. pp. 376–392. https://doi.org/10.1007/978-3-030-49570-1_26
- Tardelli, S., Avvenuti, M., Tesconi, M., & Cresci, S., 2022. Detecting inorganic financial campaigns on Twitter. *Information Systems* 103, 101769. <https://www.sciencedirect.com/science/article/pii/S0306437921000296>, <https://doi.org/10.1016/j.is.2021.101769>
- Tiwari, D., Bhati, B.S., Nagpal, B., Al-Rasheed, A., Getahun, M., & Soufiene, B. O. (2025). A swarm-optimization based fusion model of sentiment analysis for cryptocurrency price prediction. *Scientific Reports* 15, 8119. <https://www.nature.com/articles/s41598-025-92563-y>, <https://doi.org/10.1038/s41598-025-92563-y>
- Wagner, C., & Vinaimont, T. (2010). Evaluating the Wisdom of Crowds. *Issues in Information Systems*, 11, 724–732. <https://scholars.cityu.edu.hk/en/publications/evaluating-the-wisdom-of-crowds>
- Wang, S., & Vergne, J. P. (2017). Buzz Factor or Innovation Potential: What Explains Cryptocurrencies' Returns? *PLOS ONE*, 12, Article e0169556. <https://doi.org/10.1371/journal.pone.0169556>
- Wei, W. C., (2018). Liquidity and market efficiency in cryptocurrencies. *Economics Letters*, 168, 21–24. <https://www.sciencedirect.com/science/article/pii/S0165176518301320>, <https://doi.org/10.1016/j.econlet.2018.04.003>
- Williams, A., Nangia, N., & Bowman, S. (2018). *A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference*. Association for Computational Linguistics, New Orleans, Louisiana. pp. 1112–1122. <https://aclanthology.org/N18-1101/>, <https://doi.org/10.18653/v1/N18-1101>
- Wilson, T., & Jessop, S. (2019). Making sense of chaos? Algos scour social media for clues to crypto moves. Reuters <https://www.reuters.com/article/technology/making-sense-of-chaos-algos-scour-social-media-for-clues-to-crypto-moves-idUSKCN1UC0GR/>. Accessed 16 June 2025.
- Xie, P., Hailiang, C., & Hu, Y. J. (2020). Signal or Noise in Social Media Discussions: The Role of Network Cohesion in Predicting the Bitcoin Market. *Journal of Management Information Systems*, 37, 933–956. <https://doi.org/10.1080/07421222.2020.1831762>
- Xie, P. (2022). The Interplay Between Investor Activity on Virtual Investment Community and the Trading Dynamics: Evidence From the Bitcoin Market. *Information Systems Frontiers*, 24, 1287–1303. <https://doi.org/10.1007/s10796-021-10130-y>
- Yang, A., & Pangastuti, A. (2016). Stock market efficiency and liquidity: The Indonesia Stock Exchange merger. *Research in International Business and Finance*, 36, 28–40. <https://doi.org/10.1016/j.ribaf.2015.09.002>
- Yasar, B., Martin, T., & Kiessling, T. (2020). An empirical test of signalling theory. *Management Research Review*, 43, 1309–1335. <https://doi.org/10.1108/MRR-08-2019-0338>
- Yaya, O. S., Ogbonna, A. E., Mudida, R., & Abu, N. (2021). Market efficiency and volatility persistence of cryptocurrency during pre- and post-crash periods of Bitcoin: Evidence based on fractional integration. *International Journal of Finance & Economics*, 26, 1318–1335. <https://doi.org/10.1002/ijfe.1851>
- Yin, W., Hay, J., & Roth, D. (2019). *Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach*. <http://arxiv.org/abs/1909.00161>, <https://doi.org/10.48550/arXiv.1909.00161>. arXiv:1909.00161 [cs].
- Youssef, M. (2022). What Drives Herding Behavior in the Cryptocurrency Market? *Journal of Behavioral Finance*, 23, 230–239. <https://doi.org/10.1080/15427560.2020.1867142>
- Yue, W., Zhang, S., & Zhang, Q. (2021). Asymmetric News Effects on Cryptocurrency Liquidity: an Event Study Perspective. *Finance Research Letters* 41, 101799. <https://linkinghub.elsevier.com/retrieve/pii/S1544612320316135>, <https://doi.org/10.1016/j.frl.2020.101799>
- Zavolokina, L., Hein, A., Carvalho, A., Schwabe, G., & Krcmar, H. (2024). Preface to the special issue on Enterprise and organizational applications of distributed ledger technologies. *Electronic Markets*, 34, 4. <https://doi.org/10.1007/s12525-023-00688-7>
- Zeiß, C., Schaschek, M., Straub, L., Tomitza, C., & Winkelmann, A. (2024). Re-intermediation of the crypto asset ecosystem by banks: An empirical study on acceptance drivers among the populace. *Electronic Markets*, 34, 37. <https://doi.org/10.1007/s12525-024-00720-4>
- Zhang, Y., Chan, S., Chu, J., & Sulieman, H. (2020). On the Market Efficiency and Liquidity of High-Frequency Cryptocurrencies in a Bull and Bear Market. *Journal of Risk and Financial Management*, 13, 8. <https://www.mdpi.com/1911-8074/13/1/8>, <https://doi.org/10.3390/jrfm13010008>