Technology, Cybersecurity, and Cryptocurrency Returns*

Da Huang† Jeffrey Yang‡

September 10, 2024

Abstract

Are cryptocurrencies mere speculative bubbles, or do their prices reflect the underlying technology? We address this question by examining flaws in the underlying technology that could lead to cryptocurrency theft and their effect on returns. Most cryptocurrencies are open-source projects hosted on GitHub with public source code, whose cybersecurity flaws are documented to be fixed later. We find that one flaw predicts a 5 basis points decrease in the coin's daily return. A portfolio that longs no-flaw coins and shorts high-flaw ones, which we term the "cybersecurity factor," earns 30 basis points daily. Our results demonstrate that cryptocurrency prices reflect the quality of its underlying technology.

Keywords: Technology, Cryptocurrency, Cybersecurity, GitHub.
JEL Classification: G10, G12, G14

*We thank Rajesh Aggarwal, Jonathan Brogaard, Mike Cooper, Atif Ellahie, Davidson Heath, Vasudha Nair, Jiacui Li, Tim Liu, Matt Ringgenberg, and Nathan Seegert for helpful suggestions. We thank Ian Chang for excellent research assistance.

†Department of Finance, D'Amore-McKim School of Business, Northeastern University, Hayden Hall, 413B, 370 Huntington Ave, Boston, MA 02115, US. E-mail: da.huang@northeastern.edu

‡Department of Finance, David Eccles School of Business, University of Utah. 1655 Campus Drive, Salt Lake City, UT 84112, US. E-mail: jeffrey.yang@eccles.utah.edu

Technology, Cybersecurity, and Cryptocurrency Returns

ABSTRACT

Are cryptocurrencies mere speculative bubbles, or do their prices reflect the underlying technology? We address this question by examining flaws in the underlying technology that could lead to cryptocurrency theft and their effect on returns. Most cryptocurrencies are open-source projects hosted on GitHub with public source code, whose cybersecurity flaws are documented to be fixed later. We find that one flaw predicts a 5 basis points decrease in the coin's daily return. A portfolio that longs no-flaw coins and shorts high-flaw ones, which we term the "cybersecurity factor," earns 30 basis points daily. Our results demonstrate that cryptocurrency prices reflect the quality of its underlying technology.

Keywords: Technology, Cryptocurrency, Cybersecurity, GitHub.
JEL Classification: G10, G12, G14

1 Introduction

Many argue that investing in the cryptocurrency (crypto) market is akin to playing roulette because cryptos do not have an intrinsic value. Therefore, they argue that the search for crypto return predictability is futile and meaningless. The countervailing argument is that cryptocurrency returns reflect the quality of the underlying blockchain technology which may become valuable in the future (Y. Liu, Tsyvinski, & Wu, 2022). How does the market incorporate the technology in crypto returns? As a digital asset, cryptocurrencies have a unique feature that most other assets do not – they can be stolen due to flaws in the blockchain technology. Indeed, $50 billion worth of crypto has been stolen by hackers who exploit flaws in the underlying technology as of 2023 (see **Figure 1**). Investors understand this cybersecurity risk and impound the quality of the technology, i.e., hack resistance, into crypto returns (Biais, Bisiere, Bouvard, Casamatta, & Menkveld, 2023; Pagnotta, 2022). In this paper, we empirically address the question of whether crypto returns are random draws or a market mechanism reflecting the underlying blockchain technology.

We begin by measuring the flaws in the technology underlying cryptocurrencies. Most cryptocurrencies are open-source projects with public source code housed on GitHub, a platform that allows developers to create, store, manage, and share their code. The timeline of the versioning history, GH Archive, has terabytes of records including the evolution of the code base and all interactions between developers with precise timestamps at the second level. An important type of these interactions, called "Issues," are initiated by anyone who

finds a problem with the project, ranging from vague documentation to bugs in the source code. Issues describe the nature of the bug and are effectively a call for help from other developers to fix the problem.

We match 1,414 coins with market capitalization above $1 million to their corresponding GitHub repositories. Using a dictionary-based approach, we identify Issues related to bugs in the source code. We then count the number of bug-related Issues for a coin each day or week to form a *Technology Flaws* measure that captures cybersecurity exposure in coins on a daily or weekly level, respectively. Unsurprisingly, Issues related to bugs in the source code, which represent problems that can lead to hacks, are rare as 98% and 94% of coin-days and coin-weeks have zero *Technology Flaws*. On the other hand, a significant portion (roughly one-third) of coins have had technology flaws during the sample period as over 458 coins have ever had positive *Technology Flaws*.

We find that the number of *Technology Flaws* predicts future coin returns. One new technology flaw in the past day and week predicts a decrease in subsequent daily and weekly returns by 5 and 10 basis points, respectively. The predictability is stronger for small coins and remains similar to the unconditional estimates for large coins. The effect also remains robust at the intensive margin – the estimates are similar if we only use observations with positive *Technology Flaws*. These results ensure that the effect is not driven by only small coins or a large fraction of zeros in the main independent variable.

To understand the magnitude of the return predictability in the cross-section, we form

a portfolio that longs coins with zero lagged *Technology Flaws* and shorts coins in the top decile based on lagged *Technology Flaws*. Coins in the short leg of the portfolio on average have 13.4 *Technology Flaws* per day. We term this portfolio the "Cybersecurity Factor" in cryptocurrencies. The daily and weekly Cybersecurity Factors earn 30 and 88 basis points per period, respectively. While we do not claim this portfolio to be a profitable trading strategy (as it ignores costly rebalancing), the economically large magnitude of its long-short returns suggests significant return predictability in the cross-section. The magnitude may seem large but it is in line with the findings in the literature. For comparison, the weekly market, size, and momentum factors developed by Y. Liu et al. (2022) earn 167, 353, and 99 basis points, respectively. In addition, the cybersecurity factor is not explained by these three factors. We double-sort coins first by size or past return and then by *Technology Flaws*, and find the long-short portfolios earn similar returns in each size or past-return tercile. We also find that the cybersecurity factor earns an alpha of roughly 60 basis points against a 3-factor model including the crypto market, size, and momentum factors.

This paper makes several contributions. First, we address the important question – whether cryptocurrencies are mere speculative bubbles or financial assets that represent something economically meaningful. The conventional wisdom is that cryptocurrencies do not have an intrinsic value and therefore are irrational bubbles with no return predictability. We challenge this view by demonstrating that flaws in the technology underlying cryptocurrencies predict crypto returns. Our results support the view that cryptocurrency returns

reflect the quality of the underlying blockchain technology.

Second, many existing studies on cryptocurrencies model a "hacking risk" in crypto returns (Athey, Parashkevov, Sarukkai, & Xia, 2016; Biais et al., 2023; Pagnotta, 2022). To the best of our knowledge, we are the first to provide direct evidence that technology flaws are priced in crypto returns with a high-frequency measure. Lastly, the methodology of our measure opens new avenues for future research in crypto characteristics and returns. To date, most studies in this category capture crypto characteristics (e.g., technological innovation) using white papers at the initial coin offering (ICO) with cross-sectional variation (Bourveau, De George, Ellahie, & Macciocchi, 2022; Dittmar & Wu, 2019; Howell, Niessner, & Yermack, 2020; Y. Liu, Sheng, & Wang, 2021). Our methodology can be implemented to capture crypto characteristics in the code base and developer interactions at a high frequency.

This paper also contributes to a growing literature in cryptocurrency return predictability (Bouri, Lau, Lucey, & Roubaud, 2019; Y. Liu & Tsyvinski, 2021; Y. Liu et al., 2022). Many studies justify crypto return predictability with the existence of fundamental value deriving from the utility of the token (Cong, Li, & Wang, 2021; Prat, Danos, & Marcassa, 2019; Sockin & Xiong, 2020). We remain agnostic on the existence of the fundamental values of cryptos and take a different approach. As long as the price of cryptocurrencies is not exactly zero (Duchin, Solomon, Tu, & Wang, 2023), their returns are predicted by the quality of the underlying technology because cybersecurity exposure affects investors' return from holding crypto assets Biais et al. (2023); Pagnotta (2022). Our approach is similar to several studies

focused on ICOs (Bourveau et al., 2022; Dittmar & Wu, 2019; Howell et al., 2020; Y. Liu et al., 2021) but with higher resolution of the quality of technology over time.

We also add to the growing literature that studies how cybersecurity risk affects capital markets. Most studies examine how cybersecurity affects firms through the productivity or reputation of the firm (Akey, Lewellen, Liskovich, & Schiller, 2023; Crosignani, Macchiavelli, & Silva, 2023; Kamiya, Kang, Kim, Milidonis, & Stulz, 2021). This effect is then reflected in returns of the stock depending on how the firm discloses cybersecurity exposures (Amir, Levi, & Livne, 2018) and how markets understand them (T. Liu, Makridis, & Galinkin, 2023). Our study has a different and much more direct mechanism – the coin itself (analogous to the stock of the firms) can be stolen due to cybersecurity exposure, affecting how much investors value the coin (Biais et al., 2023; Pagnotta, 2022). Our Cybersecurity Factor has an important distinction with the cybersecurity risk premium in the literature. Florackis, Louca, Michaely, and Weber (2023) and Jiang, Khanna, Yang, and Zhou (2024) document that firms with higher cybersecurity risk earn higher returns. The positive risk premium suggests that the cybersecurity risk captured by the textual analysis in their studies is instantly fully priced in the market. We find that technology flaws in cryptocurrencies' code base predict a lower future return. Our results suggest that the market does not fully price the cybersecurity exposure in crypto returns, consistent with the empirical observations in T. Liu et al. (2023).

2 Data and Variables

We describe the data and variable construction in this section. We first provide an overview of GitHub, a developer platform that houses repositories (repos) for open-source projects including most cryptocurrencies. The computer code underlying these cryptocurrencies' technology is stored in GitHub repos. We then explain our methodology to capture technology flaws in cryptocurrency GitHub repos that could lead to cybersecurity exposures. Lastly, we describe the crypto trading and return data and the sample construction process.

2.1 Open Source Projects and GitHub

Most cryptocurrencies are open-source projects with code repositories housed on GitHub (Reibel, Yousaf, & Meiklejohn, 2018), a platform that stores computer code and facilitates the ongoing development of projects. GitHub is popular among developers due to its transparent versioning system. All changes to the code and interactions between developers since the project's inception are documented systematically and are viewable to the public.

The code of a project, for a cryptocurrency or not, changes and improves to add new features or fix bugs. This process can be initiated by the developers or the users. In either case, the common practice is to open an "*Issue*" which documents the features to be added or bugs to be fixed. An Issue is effectively a call for help from all coders on GitHub to work on the feature upgrade or fix the bug. Subsequently, a "*Pull Request*" is created when someone proposes a change to the code, though not all Pull Requests are accompanied by a preceding

Issue. Pull Requests are reviewed by project developers who have "admin privilege" and can approve or reject the proposed change. Approving a Pull Request creates a "*Push Event*" which implements the change into the project's code base.[1]

2.2 Measuring Technological Flaws

We construct a measure of *Technology Flaws* that captures cryptocurrencies' cybersecurity exposure due to bugs in the underlying code. We obtain *Technology Flaws* by counting the number of Issues that call for help to fix bugs. Sometimes, Issues ask for feature updates or report minor problems such as vague documentation. These Issues are not expected to lead to cybersecurity exposures. Therefore, we apply a dictionary-based filter and only count Issues that contain keywords that indicate cybersecurity exposures.[2] We create the *Technology Flaws* measure at coin-day and coin-week levels.

Table 1 presents the summary statistics of the *Technology Flaws* measure. We see that on average 0.04 and 0.27 bugs are reported for a given coin-day and coin-week, respectively. *Technology Flaws* is right-skewed. Unsurprisingly, Issues reporting bugs are rare, as 98% of coin-days and 94% of coin-weeks have zero *Technology Flaws*. On the other hand, a

[1] We only describe a simplified process with limited types of GitHub events relevant to our study. We refer readers to GitHub's REST API documentation for a detailed description of the workflow and full set of event types at https://docs.github.com/en/rest/using-the-rest-api/github-event-types.

[2] The keywords are bug, wrong, mistake, problem, incorrect, issue, bad, flaw, risk, break, breach, compromise, attack, malware, phishing, ransomware, hijack, intrusion, backdoor, and tamper. We iterate over different dictionaries that contain either broad keywords such as "bug" and "problem" or specific keywords such as "malware, phishing, ransomware, hijack, and backdoor". All empirical results remain qualitatively similar.

significant portion (roughly one-third) of coins have had technology flaws during the sample period as 458 coins have ever had positive *Technology Flaws*.

Insert **Table 1** About Here

2.2.1 Examples of GitHub Issues and Cybersecurity Exposure

We provide several concrete examples of Issues to demonstrate how technology flaws can lead to cybersecurity vulnerabilities. The following is an Issue initiated by username TippyFlitsUK on the filecoin-project project's lotus repository on November 6, 2020, at 22:59:19 GMT:

> *"I have lost a large amount of money this week primarily as a result of snapshot sync irregularities and deal completion failures. We are very much still in the early stages of Filecoin mainnet and there will obviously be bugs and obstacles to navigate. As early miners we fully accept this but it is also crucial to remember that we are now dealing with real money and livelihoods at a time of great economic instability. I for one am sincerely grateful to the developer team for their considerable efforts and excellent communication in these early days but the fact remains that bugs are an inevitability. I would respectfully like to propose that a*

contingency fund be established that will allow miners to apply for reimbursement in the unfortunate event of unavoidable financial loss."[3]

This example shows how the crypto community uses Issues as a means to identify problems and call for help. It also highlights that technology flaws can lead to monetary losses.

Issues sometimes provide specific details about bugs in the code even for prominent coins. The following is an Issue initiated by username nopara73 on the Bitcoin project's Bitcoin repository on October 2, 2018, at 00:34:21 GMT:

"At the time of writing approximately 86.51% of the Bitcoin network is vulnerable to CVE-20181-7144. It has been two weeks since 0.16.3 was tagged with the vulnerability fix. I believe the occasional breaking of user space is one of the contributing factor to this situation where things can be improved. I believe marking an RPC method deprecated and removing it within a few new release may not be sufficient, since most developers do not code directly against Bitcoin Core, rather I they take an RPC client implementation, written by someone else in their respective programming language. These projects tend to be unmaintained, incomplete, sometimes even buggy. These developers, in such an environment may consider upgrading their Core nodes risky. I know it is hard living together with legacy code and immature API design decisions, but a "never break the user space" policy may result in a more secure and stable environment."[4]

[3]See the original post at https://github.com/filecoin-project/lotus/issues/4759.
[4]See the original post at https://github.com/bitcoin/bitcoin/issues/14363.

These examples demonstrate how GitHub Issues capture flaws in cryptocurrencies' underlying technology and how *Technology Flaws* can lead to cybersecurity exposure and monetary losses.

2.2.2 Discussion of GitHub Issues and Cybersecurity Exposure

We address two questions concerning the validity of using GitHub Issues to measure technology flaws leading to cybersecurity exposures.

First, how much should we trust the quality of these "bug reports" in GitHub Issues? Are these bug reports really informative considering that people who report the bugs are not the ones who wrote the code? There are two pieces of evidence that these Issues systematically capture technology flaws. First, relying on users to report bugs is a crucial way to improve software quality and security widely used by the industry. For example, Ethereum, the second largest cryptocurrency by market capitalization as of 2023, and Coinbase, the largest US-based crypto exchange as of 2023, have bug bounty programs that incentivize outsiders to find and report bugs.[5] Second, even though it is impossible to link every Issue with the corresponding subsequent Pull Request that fixes the bug, we observe a positive correlation between the number of Issues and future Pull Requests. These observations support that GitHub Issues systematically report bugs in the code underlying cryptocurrencies.

[5]Ethereum's bug bounty program rewards up to $250,000 and can be viewed at https://ethereum.org/en/bug-bounty/. Coinbase's bug bounty program can be viewed at https://hackerone.com/coinbase. See additional background on the inception of Coinbase's bounty program based on outsiders' bug reports at https://www.coinbase.com/blog/retrospective-recent-coinbase-bug-bounty-award.

Second, do these bugs lead to serious cybersecurity failures, or do they cause mere inconvenience for user experience? Some bugs do lead to serious cybersecurity failures and, eventually, massive monetary loss. For example, in the Wormhole Network exploit, the bug allowed the hacker to "mint" 120,000 wrapped ETH (worth $325 million) out of thin air, while the bug was identified and fixed in the code base but did not go live before the exploit happened.[6] More importantly, it is difficult to assess the severity of most bugs. A bug that looks innocuous can have unexpected grave consequences. For example, an NFT project accidentally locked $34 million worth of ETH in the smart contract simply because the code did not properly add index numbers to the assets in the smart contract, which would seem almost trivial ex-ante.[7] Therefore, the bugs reported in GitHub Issues do lead to serious cybersecurity exposures beyond mere inconvenience for user experience.

To summarize, since GitHub Issues report bugs that could lead to cybersecurity exposures and monetary losses, we develop the measure of technology flaws using these Issues and examine whether it predicts crypto returns in subsequent analyses.

[6]See the Solana blockchain record of the hacker "minting" 120,000 wETH at https://solscan.io/tx/2zCz2GgSoSS68eNJENWrYB48dMM1zmH8SZkgYneVDv2G4gRsVfwu5rNXtK5BKFxn7fSqX9BvrBc1rdPAeBEcD6Es. See the bug fix in Wormhole Network's code base that would have prevented the exploit at https://github.com/wormhole-foundation/wormhole/commit/7edbbd3677ee6ca681be8722a607bc576a3912c8#diff-0d27d8889edd071b86d3f3299276882d97613ad6ab3b0b6412ae4ebf3ccd6370L92-R101. See more details about this exploit at https://www.theverge.com/2022/2/3/22916111/wormhole-hack-github-error-325-million-theft-ethereum-solana and https://github.com/liqtags/crypto-rekts/blob/main/rekts/Wormhole.md.

[7]See more for details about the incident of the Aku Ethereum NFT launch at https://decrypt.co/98530/aku-ethereum-nft-launch-ends-with-34m-locked-in-flawed-smart-contract.

2.3 Crypto Trading Data and Sample Construction

Finally, we obtain daily and weekly crypto trading data, including trading volume, market capitalization, and price, from CoinGecko for 14,909 coins. We require that a coin enters the sample after having a $1 million market capitalization and never leaves the sample afterward unless its trading stops altogether. We then manually match coins to their GitHub repository and discard ones that we could not match. The final sample consists of 1,414 cryptocurrencies from 2014 to 2023. Summary statistics for daily and weekly returns, trading volumes, and market capitalizations are presented in **Table 1**. The average daily and weekly returns for the sample coins are 0.36% and 1.73%, respectively. The market cap of a coin can fall below $1 million and stay in the sample after falling below the initial $1 million threshold to enter.

3 Technological Flaws and Crypto Returns

In this section, we examine the relation between *Technology Flaws* and future crypto returns. We show that *Technology Flaws* negatively predict subsequent crypto returns and that the negative predictability is robust.

3.1 Technology Flaws and Crypto Returns

We examine whether *Technology Flaws* predict future crypto returns with the following equation:

$$\text{Return}_{i,t} = \beta_0 + \beta_1 \, \textit{Technology Flaws}_{i,t-1} + \boldsymbol{\beta} \boldsymbol{X}_{i,t} + \zeta_t + \phi_i + \epsilon_{i,t}, \tag{1}$$

where i indexes coin and t indexes day or week. $\boldsymbol{X}_{i,t}$ is a vector of control variables including the log of trading volume and the log of market cap. ζ_t represents fixed effects on daily or weekly levels that sweep out factors that apply to all coins in the market such as the overall sentiment in the crypto market. ϕ_i is coin fixed effects that sweep out coin-specific time-invariant characteristics such as how prominent a coin is. Returns are expressed in percentage points and *Technology Flaws* is expressed as a count. **Table 2** reports the results.

Insert **Table 2** About Here

Table 2 shows that *Technology Flaws* negatively predict future crypto returns. Panels A and B report daily and weekly results, respectively. In both panels, columns (1) to (4) show specifications with different fixed effects. Column (1) estimates **Equation 1** without any fixed effects. Columns (2) and (3) add time and coin fixed effects, respectively. Column (4) adds both time and coin fixed effects. We see that *Technology Flaws* negatively predicts

future crypto returns consistently across all specifications with statistical significance at the 1% level and a similar economic magnitude. We focus on the most saturated specification with the most conservative estimates in column (4). If one additional Issue reporting bugs in the crypto code is posted on GitHub, the corresponding coin sees a decrease of 5 (10) basis points in the return over the next day (week). The effect of *Technology Flaws* on crypto returns is economically large – an additional *Technology Flaw* reduces the return in the next day (week) by 19% (6%) of the unconditional mean return in the sample.

3.2 Robustness of the Return Predictability

We address two concerns about the negative predictability of *Technology Flaws* on crypto returns. First, we show that the predictability is not just concentrated in small coins. A large empirical asset pricing literature suggests that return predictability is more likely to be concentrated in illiquid assets with small market caps (Asparouhova, Bessembinder, & Kalcheva, 2010; Chu, Hirshleifer, & Ma, 2020). We examine whether *Technology Flaws* predict returns only for small coins by estimating Equation 1 separately for small and large terciles based on lagged market caps. **Table 3** reports the result.

Insert **Table 3** About Here

We see that the return predictability is stronger in small coins but remains robust in large

coins. Columns (1) to (4) show the estimates for small coins. On average, these estimates are three to five times larger compared to the baseline results in **Table 2**, except for column (4) in the weekly panel. Columns (5) to (8) show the estimates for large coins. The point estimates are similar to those in the baseline results and remain statistically significant. As a result, we conclude that the negative return predictability is not entirely driven by small coins and remains robust with large coins.

Second, we show that the results are not driven by a large number of zeros in the independent variable, *Technology Flaws*. Since over 98% of coin-days and 94% of coin-weeks have *Technology Flaws* of zero, it is possible that the extremely right-skewed distribution gives rise to spurious results. To address this concern, we estimate Equation 1 only with observations with positive *Technology Flaws*. This estimate allows us to obtain the intensive margin of the effect.

Insert **Table 4** About Here

Table 4 shows that the results remain qualitatively similar with positive *Technology Flaws* observations. The results alleviate the concern that the predictability is driven by the right-skewed distribution with a large fraction of zeros in *Technology Flaws*.

4 Cybersecurity Factor

In this section, we examine the cross-sectional predictability of *Technology Flaws* on returns by forming a long-short portfolio based on *Technology Flaws*. We term this portfolio the "Cybersecurity Factor." We show that the Cybersecurity Factor is profitable and robust to double sorting against size and momentum. Finally, we show that the Cybersecurity Factor earns a meaningful alpha against existing factors in the cryptocurrency return predictability literature.

4.1 Cybersecurity Factor Returns

Following Y. Liu et al. (2022), we develop a Cybersecurity Factor at both daily and weekly levels. In each period, the portfolio longs coins with zero lagged *Technology Flaws* and shorts coins with lagged *Technology Flaws* in the top decile.[8] On average, coins in the short leg have 13.4 *Technology Flaws* **Table 5** shows the average return of the Cybersecurity Factor on daily and weekly levels.

Insert **Table 5** About Here

Table 5 shows that the Cybersecurity Factor generates significant returns. On average,

[8]When the 90th percentile of *Technology Flaws* is zero for a day or week, there are less than 10% of coins with positive *Technology Flaws* during that period. We handle this situation in two ways. First, we simply skip days and weeks when it happens. Second, we short coins with zero lagged *Technology Flaws* and long coins with positive lagged *Technology Flaws*. The results are qualitatively similar.

the long-short portfolio earns a daily and weekly return of 30 bps and 88 bps, respectively. These returns are statistically and economically significant. Holding the Cybersecurity Factor and rebalancing weekly from 2017 to 2023 earns a cumulative return of roughly 297%. This number may seem high but is in line with the literature. We plot the Cybersecurity Factor along with the crypto market, size, and momentum factors developed by Y. Liu et al. (2022) in Figure 2. The crypto market, size, and momentum factors earn a cumulative return of roughly 544%, 1,149%, and 323% over the same period. Similar to the argument in Y. Liu et al. (2022), while the Cybersecurity Factor is highly profitable on paper (with a Sharpe ratio of 2.28 and 1.20 on the daily and weekly levels, respectively), it does not account for transaction costs from frequent rebalancing. But importantly, the high profitability on paper suggests a strong cross-sectional return predictability of *Technology Flaws*.

Insert **Figure 2** About Here

4.2 Double Sorts on Size and Momentum

A potential concern is that the Cybersecurity Factor may be effectively sorting on coin size if size and *Technology Flaws* are correlated. In addition, it may also sort on past returns if there is any persistence in cryptocurrency returns. To address these concerns, we double-sort coins into size or past return terciles and then by *Technology Flaws* on the weekly level.

Table 6 presents the results. Panels A and B report returns of the portfolios double-sorted first on size and past returns, respectively.

Insert **Table 6** About Here

We see that the Cybersecurity Factor remains robust after conditioning on size and past returns. Panel A shows that the cross-sectional predictability is stronger in large coins. The long-short return is statistically insignificant for small coins but remains economically large. Panel B shows that the long-short return is higher among coins with low past returns. Similarly, even though the return is statistically insignificant for coins with high past returns, it remains economically significant.

4.3 Cybersecurity Factor and Existing Factors

Lastly, we explicitly contrast our Cybersecurity Factor against the existing crypto factors of market, size, and momentum (Y. Liu et al., 2022). We estimate time series regressions on the weekly level to obtain the alpha of the Cybersecurity Factor against existing factors.[9] The results are presented in **Table 7**.

Insert **Table 7** About Here

[9]We thank Yukun Liu for making the crypto factors available on his website. Due to data availability, the sample period is from 2017 to 2023.

We see that the cybersecurity factor is not subsumed by the existing factors and earns a statistically and economically significant alpha. The results show that our long-short portfolio captures the cybersecurity exposure stemming from technology flaws that are not explained by known cryptocurrency factors.

5 Conclusion

In this paper, we provide empirical evidence that cryptocurrency returns are not mere random draws, but instead reflect the underlying quality of the blockchain technology. By constructing a high-frequency measure of *Technology Flaws* from GitHub data, we show that these flaws in cryptocurrencies' code bases related to cybersecurity vulnerabilities predict lower future returns for cryptocurrencies. Our findings challenge the prevailing narrative that cryptocurrencies have unpredictable and completely random returns due to their lack of intrinsic values. The Cybersecurity Factor demonstrates the importance of technological quality and cybersecurity exposure as key drivers of returns of digital assets.

References

Akey, P., Lewellen, S., Liskovich, I., & Schiller, C. (2023). Hacking corporate reputations. *Working Paper*.

Amir, E., Levi, S., & Livne, T. (2018). Do firms underreport information on cyber-attacks? evidence from capital markets. *Review of Accounting Studies*, *23*, 1177–1206.

Asparouhova, E., Bessembinder, H., & Kalcheva, I. (2010). Liquidity biases in asset pricing tests. *Journal of Financial Economics*, *96*(2), 215–237.

Athey, S., Parashkevov, I., Sarukkai, V., & Xia, J. (2016). Bitcoin pricing, adoption, and usage: Theory and evidence. *Working paper*.

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

Bouri, E., Lau, C. K. M., Lucey, B., & Roubaud, D. (2019). Trading volume and the predictability of return and volatility in the cryptocurrency market. *Finance Research Letters*, *29*, 340–346.

Bourveau, T., De George, E. T., Ellahie, A., & Macciocchi, D. (2022). The role of disclosure and information intermediaries in an unregulated capital market: evidence from initial coin offerings. *Journal of Accounting Research*, *60*(1), 129–167.

Chu, Y., Hirshleifer, D., & Ma, L. (2020). The causal effect of limits to arbitrage on asset pricing anomalies. *Journal of Finance*, *75*(5), 2631–2672.

Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic adoption and valuation.

Review of Financial Studies, *34*(3), 1105–1155.

Crosignani, M., Macchiavelli, M., & Silva, A. F. (2023). Pirates without borders: The propagation of cyberattacks through firms' supply chains. *Journal of Financial Economics*, *147*(2), 432–448.

Dittmar, R. F., & Wu, D. A. (2019). Initial coin offerings hyped and dehyped: An empirical examination. *Working paper*.

Duchin, R., Solomon, D., Tu, J., & Wang, X. (2023). The cryptocurrency elephant in the room. *Working Paper*.

Florackis, C., Louca, C., Michaely, R., & Weber, M. (2023). Cybersecurity risk. *Review of Financial Studies*, *36*(1), 351–407.

Howell, S. T., Niessner, M., & Yermack, D. (2020). Initial coin offerings: Financing growth with cryptocurrency token sales. *Review of Financial Studies*, *33*(9), 3925–3974.

Jiang, H., Khanna, N., Yang, Q., & Zhou, J. (2024). The cyber risk premium. *Management Science*.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A., & Stulz, R. M. (2021). Risk management, firm reputation, and the impact of successful cyberattacks on target firms. *Journal of Financial Economics*, *139*(3), 719–749.

Liu, T., Makridis, C., & Galinkin, E. (2023). Do financial markets understand cybersecurity: Evidence from a direct measure of firm network vulnerabilities. *Working Paper*.

Liu, Y., Sheng, J., & Wang, W. (2021). Technology and cryptocurrency valuation: Evidence

from machine learning. *Working paper*.

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

Liu, Y., Tsyvinski, A., & Wu, X. (2022). Common risk factors in cryptocurrency. *Journal of Finance*, *77*(2), 1133–1177.

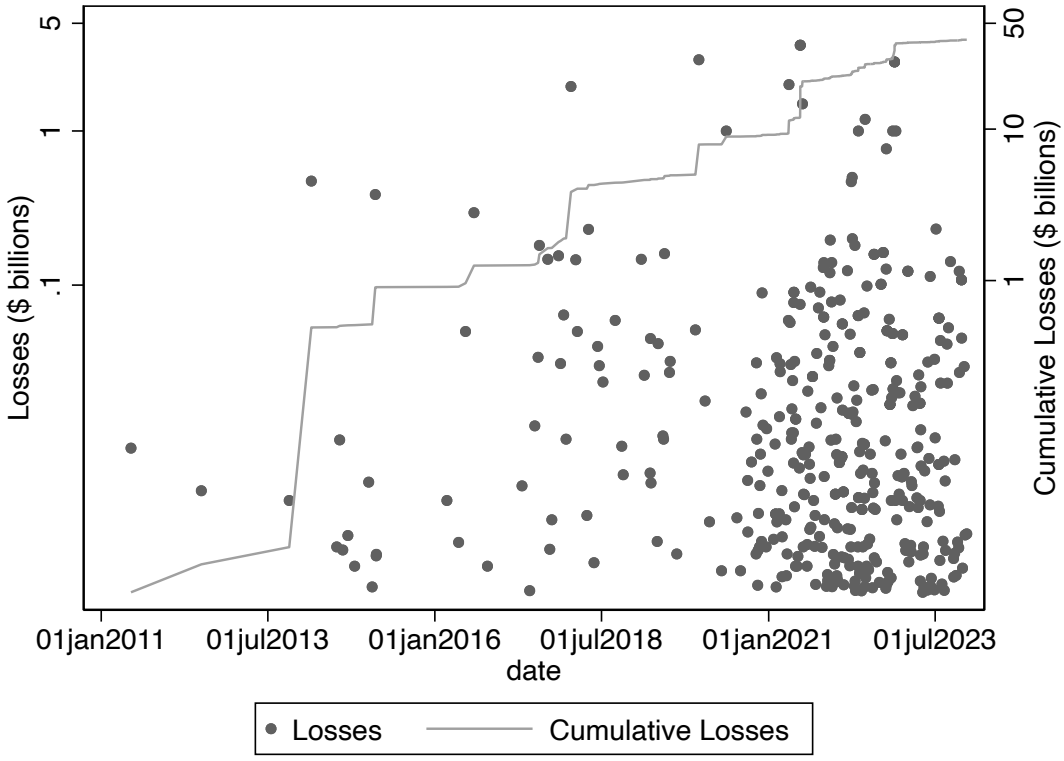
Pagnotta, E. S. (2022). Decentralizing money: Bitcoin prices and blockchain security. *Review of Financial Studies*, *35*(2), 866–907.

Prat, J., Danos, V., & Marcassa, S. (2019). Fundamental pricing of utility tokens. *Working paper*.

Reibel, P., Yousaf, H., & Meiklejohn, S. (2018). Why is a Ravencoin like a Tokendesk? An exploration of code diversity in the cryptocurrency landscape. *Working Paper*.

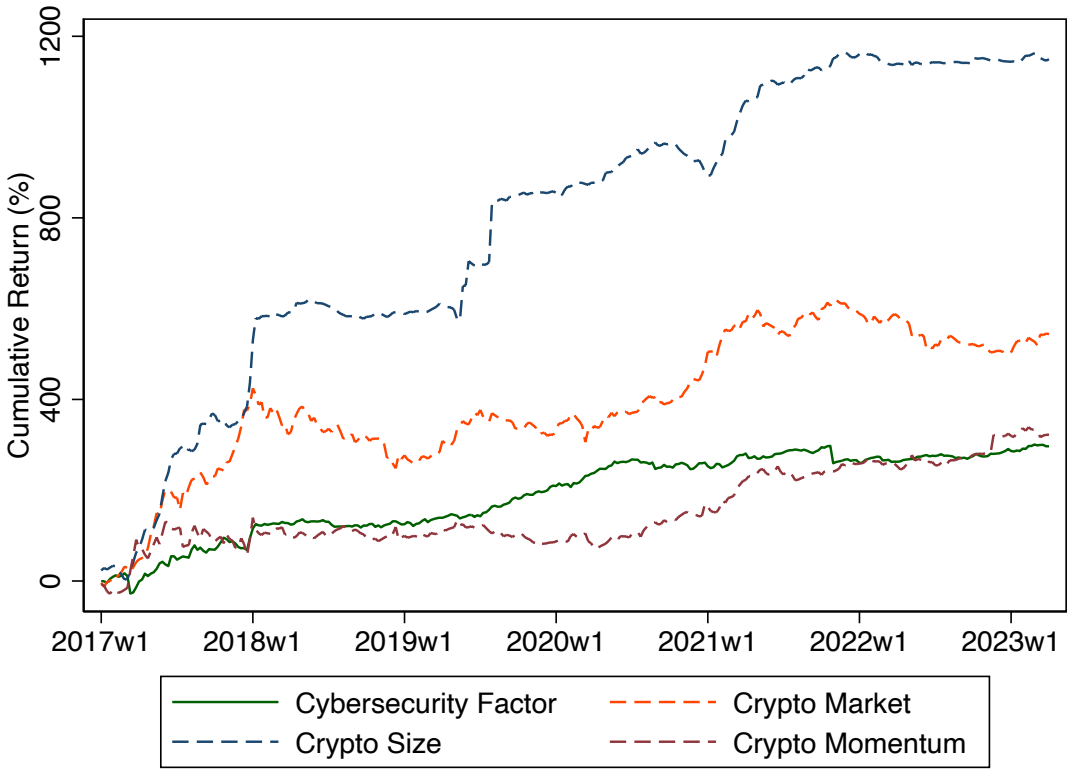
Sockin, M., & Xiong, W. (2020). A model of cryptocurrencies. *Working paper*.

Figure 1. Cryptocurrency Thefts Over Time



The figure plots crypto hacks over time. Each dot corresponds to a hack, with the x-axis representing the date of the event and the y-axis showing the associated monetary loss in billions of dollars. The solid line plots the cumulative losses from these hacks. The data is collected from an open-source repository that is dedicated to recording cryptocurrency thefts (see https://github.com/liqtags/crypto-rekts).

Figure 2. Cybersecurity Factor Cumulative Return



The figure plots the cumulative return (without compounding) of the cybersecurity factor along with the crypto market, size, and momentum factors from Y. Liu et al. (2022).

Table 1
Summary Statistics

This table presents summary statistics of key variables at the daily and weekly levels. *Technology Flaws* is the count of issue events that contain at least one of the keywords specified in the content filter. Returns are the daily or weekly percentage returns for crypto i at time t. Trading volume and market cap are expressed in millions of dollars for crypto i at time t. The sample contains 1,414 unique coins from 2013 to 2023.

	Mean	StdDev	p10	Median	p90	N
Daily						
Technology Flaws (#)	0.04	0.40	0.00	0.00	0.00	1,605,737
Returns (%)	0.36	10.43	-8.70	-0.15	8.94	1,605,737
Volume ($ millions)	21.76	111.12	0.00	0.13	14.65	1,605,737
Market Cap ($ millions)	200.54	920.03	0.10	5.98	211.48	1,605,588
Weekly						
Technology Flaws (#)	0.27	2.10	0.00	0.00	0.00	231,568
Returns (%)	1.73	26.14	-21.55	-1.33	24.65	230,146
Volume ($ millions)	150.99	761.69	0.00	0.97	110.57	231,568
Market Cap ($ millions)	198.67	915.76	0.09	5.79	208.48	231,558

Table 2
Technological Flaws and Cryptocurrency Returns

The table presents the results of the following regression:

$$\text{Return}_{i,t} = \beta_0 + \beta_1 \text{Technology Flaws}_{i,t-1} + \boldsymbol{\beta} \boldsymbol{X}_{i,t} + \phi_i + \zeta_t + \epsilon_{i,t},$$

where $\text{Return}_{i,t}$ is the return for crypto i on date t in percentage points, $\text{Technology Flaws}_{i,t}$ is the number of bug reports for crypto i on date t, and $\boldsymbol{X}_{i,t}$ is a vector of control variables including the log of trading volume and the log of market cap. ζ and ϕ are time and coin fixed effects, respectively. In the daily (weekly) specification, ζ represents date (week) fixed effects. Panel A presents daily results, while Panel B presents weekly results. Robust standard errors clustered at the coin level are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Daily

	(1)	(2)	(3)	(4)
Technology Flaws	-0.07***	-0.07***	-0.09***	-0.05***
	(0.02)	(0.02)	(0.02)	(0.02)
ln(Market Cap)	0.02	0.01	0.10***	0.24***
	(0.01)	(0.02)	(0.03)	(0.03)
ln(Volume)	0.01	0.01	0.15***	0.17***
	(0.01)	(0.01)	(0.02)	(0.02)
Coin FE	No	No	Yes	Yes
Date FE	No	Yes	No	Yes
Adjusted R-squared	0.000	0.124	0.003	0.128
Observations	1,492,295	1,492,295	1,492,290	1,492,290

Panel B: Weekly

	(1)	(2)	(3)	(4)
Technology Flaws	-0.10***	-0.09***	-0.18***	-0.10***
	(0.03)	(0.02)	(0.04)	(0.03)
ln(Market Cap)	0.63***	0.57***	1.32***	1.61***
	(0.06)	(0.06)	(0.13)	(0.17)
ln(Volume)	-0.24***	-0.25***	0.08	0.15**
	(0.05)	(0.05)	(0.07)	(0.07)
Coin FE	No	No	Yes	Yes
Week FE	No	Yes	No	Yes
Adjusted R-squared	0.002	0.180	0.008	0.186
Observations	214,558	214,558	214,548	214,548

Table 3
Technological Flaws and Cryptocurrency Returns by Coin Size

The table repeats the analyses in **Table 2** but estimates the effect in subsamples of terciles based on the lagged coin market cap. Robust standard errors clustered at the coin level are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Daily

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Small				Large			
Technology Flaws	-0.39***	-0.36**	-0.34*	-0.21	-0.09***	-0.08***	-0.10***	-0.03*
	(0.11)	(0.15)	(0.18)	(0.17)	(0.02)	(0.02)	(0.02)	(0.02)
Coin FE	No	No	Yes	Yes	No	No	Yes	Yes
Date FE	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.004	0.067	0.010	0.078	0.004	0.302	0.009	0.310
Observations	494,967	494,967	494,948	494,948	498,135	498,134	498,112	498,111

Panel B: Weekly

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Small				Large			
Technology Flaws	-0.31**	-0.19*	-0.44***	-0.01	-0.11***	-0.12***	-0.19***	-0.08**
	(0.12)	(0.11)	(0.17)	(0.18)	(0.03)	(0.02)	(0.04)	(0.04)
Coin FE	No	No	Yes	Yes	No	No	Yes	Yes
Week FE	No	Yes	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-squared	0.014	0.129	0.028	0.147	0.009	0.326	0.019	0.344
Observations	71,636	71,636	71,596	71,596	71,333	71,333	71,292	71,292

Table 4
Technological Flaws and Cryptocurrency Returns with Positive Technology Flaws

The table repeats the analyses in **Table 2** but use only coin-days or coin-weeks with positive $TechnologyFlaws_{t-1}$ for estimation. Robust standard errors clustered at the coin level are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Daily

	(1)	(2)	(3)	(4)
Technology Flaws	-0.05***	-0.02	-0.12***	-0.08***
	(0.02)	(0.01)	(0.02)	(0.03)
Controls	Yes	Yes	Yes	Yes
Coin FE	No	No	Yes	Yes
Date FE	No	Yes	No	Yes
Adjusted R-squared	0.002	0.306	0.013	0.321
Observations	29,669	29,642	29,571	29,544

Panel B: Weekly

	(1)	(2)	(3)	(4)
Technology Flaws	-0.04	-0.03**	-0.14***	-0.11***
	(0.02)	(0.02)	(0.05)	(0.04)
Controls	Yes	Yes	Yes	Yes
Coin FE	No	No	Yes	Yes
Week FE	No	Yes	No	Yes
Adjusted R-squared	0.003	0.321	0.009	0.330
Observations	13,233	13,233	13,134	13,134

Table 5
Cybersecurity Factor Performance

This table presents the average portfolio returns for the long, short, and long-short portfolios. The long portfolio comprises of coins that had no bugs (*Technology Flaws* = 0) in period $t-1$, while the short portfolio is comprised of coins that were in the top 10th percentile of *Technology Flaws* in period $t-1$. We exclude dates where all coins had *Technology Flaws* = 0 in period $t-1$. All portfolios are equal-weighted. Portfolio returns are expressed in percentages and robust standard errors are reported in parentheses. Panel A presents daily results, while Panel B presents weekly results.

Panel A: Daily Factor

	Long	Short	Long - Short
Return	0.55***	0.25**	0.30***
	(0.08)	(0.10)	(0.05)
N	2541	2528	2528

Panel B: Weekly Factor

	Long	Short	Long - Short
Return	3.08***	2.20***	0.88***
	(0.68)	(0.73)	(0.28)
N	364	363	363

Table 6
Cybersecurity Factor Performance by Momentum and Size Tercile

This table presents the average portfolio returns for the long, short, and long-short cybersecurity portfolios, double sorted against momentum and size terciles. We initially sort into momentum or size terciles. Momentum is sorted based on the second lag of weekly returns, while size is sorted based on lagged market cap. For each size or momentum tercile, we then compute the long-short Cybersecurity portfolio. The long portfolio is comprised of coins that had no bugs (*Technology Flaws* $= 0$) in period $t-1$, while the short portfolio is comprised of coins that were in the top 10th percentile of *Technology Flaws* in period $t-1$. We exclude tercile-dates where all coins had *Technology Flaws* $= 0$ in period $t-1$. All portfolios are equal-weighted. Portfolio returns are expressed in percentages and robust standard errors are reported in parentheses. Panel A presents daily results, while Panel B presents weekly results.

Panel A: Double-Sort, Size then *Technology Flaws*

	Long	Short	Long - Short
Small Market Cap	1.83***	1.04	0.59
	(0.63)	(1.32)	(1.09)
	2.59***	0.76	1.51**
	(0.71)	(0.77)	(0.57)
Large Market Cap	3.58***	2.48***	1.10**
	(0.76)	(0.73)	(0.37)

Panel B: Double-Sort, Momentum then *Technology Flaws*

	Long	Short	Long - Short
Low Past Return	3.24***	1.87**	1.34**
	(0.70)	(0.87)	(0.55)
	2.42***	1.55**	0.85**
	(0.67)	(0.75)	(0.36)
High Past Return	3.51***	2.99***	0.49
	(0.70)	(0.87)	(0.55)

Table 7
Cybersecurity Factor and Crypto Market, Size, and Momentum Factors

The table presents the results of the following regression:

$$\text{CybersecurityFactor}_t = \alpha + \beta_1 \text{CryptoMarket}_t + \beta_2 \text{CryptoSize}_t + \beta_3 \text{CryptoMomentum}_t + \epsilon_t,$$

where $\text{CybersecurityFactor}_t$ is the return for the cybersecurity factor in week t in percentage points. CryptoMarket_t, CryptoSize_t, and CryptoMomentum_t are the Y. Liu et al. (2022) crypto market, size, and momentum factors, respectively. Newey West standard errors are estimated with one lag and reported in parentheses.

	(1)	(2)	(3)	(4)
Crypto Market	-0.03			-0.05*
	(0.03)			(0.03)
Crypto Size		0.10**		0.11***
		(0.04)		(0.04)
Crypto Momentum			0.04	0.02
			(0.08)	(0.07)
α	0.96***	0.56**	0.87***	0.60**
	(0.31)	(0.27)	(0.29)	(0.26)
Observations	325	325	325	325