Hornuf, Lars; Momtaz, Paul P.; Nam, Rachel J.; Yuan, Ye

**Working Paper**

# Cybercrime on the Ethereum Blockchain

CESifo Working Paper, No. 10598

**Provided in Cooperation with:**
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Hornuf, Lars; Momtaz, Paul P.; Nam, Rachel J.; Yuan, Ye (2023) : Cybercrime on the Ethereum Blockchain, CESifo Working Paper, No. 10598, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/279349

CESifo WORKING PAPERS

10598
2023

August 2023

# Cybercrime on the Ethereum Blockchain

*Lars Hornuf, Paul P. Momtaz, Rachel J. Nam, Ye Yuan*

CESifo

## Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10598*

# Cybercrime on the Ethereum Blockchain

## Abstract

We propose a taxonomy of cybercrime on the Ethereum blockchain and examine how cybercrime impacts victims' risk-taking and returns. Our difference-in-differences analysis of a sample of victims and matched non-victims suggests that victims increase their long-term total risk-taking and earn lower risk-adjusted returns in the post-cybercrime period. Victims' long-term total risk-taking increases because they increase diversifiable risk in the long term. The increased diversifiable risk correlates with victims' withdrawal from altcoins after cybercrime. At the same time, the reduction in risk-adjusted returns correlates with increased trading activity and churn, due plausibly to managing cybercrime exposure. In the cross-section of Ethereum addresses, we show that the most-affluent victims take a systematic approach to restore their pre-cybercrime wealth level, while the least-affluent victims turn into gamblers. Finally, a parsimonious forensic model explains a good part of the addresses' probability of being involved in cybercrime, both on the victim and the cybercriminal side.

JEL-Codes: G140, G240, G300, L260, M130, O160.

Keywords: Ethereum blockchain, market manipulation, financial fraud, token investment scam, cybercrime, cryptocurrency.

*Lars Hornuf*
*Chair of Business Administration*
*Technical University of Dresden / Germany*
*lars.hornuf@tu-dresden.de*

*Paul P. Momtaz**
*TUM School of Management*
*Technical University of Munich / Germany*
*momtaz@tum.de*

*Rachel J. Nam*
*Goethe-University Frankfurt / Germany*
*rachel.nam@outlook.com*

*Ye Yuan*
*Technical University of Munich / Germany*
*yuanyuan125@icloud.com*

*corresponding author

This article evolved as part of the research project 'Cybercrime on the Ethereum Blockchain', which has been supported by the Frankfurter Institut für Risikomanagement und Regulierung (FIRM). We thank seminar participants at the 3rd Boca Corporate Finance and Governance Conference at Florida Atlantic University and at the Diginomics Seminar at the University of Bremen.

# 1 Introduction

A substantial share of financial market misconduct and fraud takes place on blockchains. The Federal Trade Commission (FTC) reports in a recent study that more than $1 out of $4 that are reported stolen was stolen in cryptocurrency. The FTC documents $1.18 billion in aggregate losses to cybercriminals since 2018, with most losses in Bitcoin (70%), Tether (10%), and Ether (9%).[1] Using several terabytes of primary blockchain data from Ethereum and the fact that these on-chain scams are readily observable on public blockchains, we evidence that the FTC underestimates crypto scams by more than an order of magnitude. Relative to the FTC's estimate of scams on the Ethereum blockchain amounting to $106 million, we are able to use publicly available primary blockchain data to show that Ethereum addresses associated with scams received a staggering $1.65 billion—almost 16 times what was reported by the FTC.

Our study is among the first to study on-chain market misconduct and fraud at an aggregate scale on the Ethereum blockchain, using primary ledger data in their entirety. More than Bitcoin, Ethereum is intriguing from a forensic perspective because its ability to host smart contracts leads to a broader range of cybercrimes. For the purpose of identifying cybercrimes on Ethereum as such, we rely on crowd-reported incidents of alleged scams on *Etherscan*, cooperating with market experts from *ScamAlert* to validate reported scams. Based on these *confirmed* scams, we develop a taxonomy of scams on the Ethereum blockchain, extending Cumming et al.'s (2021) and Hornuf et al.'s (2022) approaches, and identify 19 unique cybercrime categories. Table 1 describes the fraud categories we were able to identify and their respective relevance. The economically most significant categories are Ponzi schemes, which make up 60% of the aggregate stolen funds on Ethereum; followed by giveaways (18%), exploits (13%), and hacks (5%).

With the taxonomy of cybercrime on the Ethereum blockchain in hand, our empirical analyses focus on how cybercrime impacts victims' address-level risk–return trade-off in a causal difference-in-differences framework. Specifically, our study aims to explore (1) how victims of cybercrime adjust their address-level risk-taking and (2) how the post-cybercrime adjustment to risk-taking levels is reflected in victims' address-level risk-adjusted returns. To this end, we implemented a pre- versus post-cybercrime comparison of victim and matched non-victim addresses with the in-

[1] https://www.ftc.gov/system/files/ftc_gov/pdf/Crypto%20Spotlight%20FINAL%20June%202022.pdf, retrieved July 25, 2023.

stant at which a cybercrime became public knowledge as the event date. Methodologically, this involves three preparatory steps. First, although we verified with external experts each individual cybercrime, we did not have the exact date at which a cybercrime was publicly identified. Thus, in order to determine the precise event timing of when a cybercrime became public knowledge, we manually researched social media for the first mention of a certain activity being a scam. Second, given the high dimensionality and imbalance of our primary ledger data, we implemented a Euclidean distance approach to pair victim addresses with matching non-victim/non-cybercriminal addresses to ensure that our difference-in-differences model correctly identifies average treatment effects. Third, for each address on the Ethereum blockchain, we estimated Liu et al.'s (2022) three-factor crypto-asset pricing model in order to characterize victim and matching non-victim addresses by risk-taking levels and risk-adjusted returns (i.e., alphas).

Figure 1 illustrates our *initial* finding that victims' *raw* returns (i.e., non-risk-adjusted returns) increase after a cybercrime. We follow Barber and Odean (2000) measuring gross monthly raw returns as the change in address-level token prices at the end of the month relative to the beginning-of-month prices for all tokens held at the beginning of the month. By implication, Barber and Odean (2000) monthly raw returns only account for the *behavioral* effect of cybercrime on returns, not for the misappropriated funds due to the cybercrime *per se*. The graph illustrates that victims become better investors post-cybercrime. However, if one were to account for the nominal value of abducted funds due to the cybercrime, cybercrime victims have, on average, lost 10% of their wealth twelve months after the cybercrime relative to matched non-victims. Second, Figure 1 illustrates nearly perfect parallel trends for the treatment (cybercrime victims) and control observations (matched non-victims), suggesting strong matching results and hence the identification of a causal effect of cybercrime on victim behavior.

**[Place Figure 1 about here.]**

Strikingly, although victims' raw returns *increase* after a cybercrime, their risk-adjusted returns *decrease* statistically and economically significantly. We regress address-level alphas from Liu et al.'s (2022) three-factor crypto-asset pricing model in a difference-in-differences model and find highly significant marginal effects, suggesting that victims' alphas respond significantly negatively to cybercrime. In terms of economic magnitude, victims' risk-adjusted returns in the post-cybercrime

period reduce by 55.2 to 96.4 percentage points relative to matched non-victims. Therefore, while cybercrime victims' raw returns respond positively to cybercrime, their risk-adjusted returns respond negatively, indicating that victims may increase their risk-taking levels after being scammed.

Consistent with our conjecture that the discrepancy between positive raw and negative risk-adjusted returns for cybercrime victims can be explained by higher post-cybercrime risk-taking, we confirm that total risk-taking increases in the long term. The average treatment effect of cybercrime on victims' total risk-taking twelve months after the event is in the 5.7 to 8.1 percentage-point range. It should be noted, however, that cybercrime victims' total risk-taking level reduces in the short term (i.e., three to six months after the cybercrime), increases in the medium term (i.e., six to twelve months after the cybercrime), and then remains permanently at a level that is higher than the initial risk-taking level in the long term (i.e., after twelve months). This result is consistent with recent literature showing that investor behavior, such as risk appetite, can change over time and that the level of risk—such as the level of risk of falling victim to cybercrime—is itself a determinant of such changes (Dicle, 2019). Further, we conduct a risk decomposition and split total address-level risk-taking into diversifiable and non-diversifiable risk-taking levels per address. Interestingly, we find that the post-cybercrime response of victims in terms of *total* risk-taking is mostly driven by changes in their *diversifiable* risk-taking, both in terms of economic magnitude and the time structure of the treatment effects (i.e., lower risk-taking in the short term and higher risk-taking in the long term). As for non-diversifiable risk-taking, we report average treatment effects of cybercrime that are constantly decreasing in the post-cybercrime period, reaching a permanent level that is between 0.8% and 4.6% percentage points lower than the initial pre-cybercrime level after twelve months. Taken together, the negative average treatment effect of cybercrime on victims' risk-adjusted returns is driven by increased diversifiable risk-taking, while non-diversifiable risk-taking reduces.

We also examine the heterogeneous responses of victims to different cybercrime categories. Post-cybercrime blockchain address-level risk-adjusted returns and risk-taking critically depends on the type of cybercrime a victim fell for. For risk-adjusted returns, fake token scams, darkweb activity, and sextortion have a positive effect on victims' post-event alphas, while Ponzi schemes, phishing scams, investment scams, hacks, and exploits have a negative effect. For total risk-taking, Ponzi schemes, events on the darkweb, and sextortion increase victims' risk-taking levels, while

giveaways, investment scams, hacks, and exploits reduce it. By far the economically most significant increase in total risk-taking occurs after *darkweb-related scams*, when victims double their total risk-taking. In contrast, the economically most significant reduction in total risk-taking occurs after *investment scams*, when victims reduce their total risk-taking by more than half.

Given the overarching result that victims of cybercrime increase total risk-taking which reduces their risk-adjusted returns, we next explore potential mechanisms that help explain the finding. The evidence from triple difference models suggests two overarching behavioral explanations for how cybercrime changes victims' risk–return trade-off. First, victims of cybercrime significantly increase their trading activity and churn rate, which loads significantly negatively on the alpha-related triple difference estimator. This suggests that higher trading activity reduces risk-adjusted returns, which is in line with the evidence by Odean and Barber (1999) for traditional finance and Sokolov (2021) for decentralized finance. Second, address-level token diversification and ownership of different token categories (including altcoins and stablecoins) load positively on risk-adjusted returns and non-diversifiable risk and negatively on diversifiable risk. Overall, the collective evidence indicates that the increase in diversifiable risk, which reduces risk-adjusted returns, is largely caused by victims divesting altcoins.

Additionally, we investigate heterogeneous treatment effects for Ethereum addresses of various wealth levels (i.e., comparing the top 10% to the bottom 10% in terms of pre-cybercrime address balance). We document that the least affluent Ethereum addresses' risk-adjusted returns decrease more than those of the most affluent addresses. Again, the discrepancy in responses of cybercrime victims of differential wealth can be explained by their responses in risk-taking levels, which is in line with Guiso and Paiella's (2008) finding that equity investors who are more likely to become liquidity-constrained exhibit a higher degree of absolute risk aversion. The least affluent addresses dramatically increase their total risk-taking relative to the most affluent addresses following a cybercrime. However, the least affluent only increase their address-level diversifiable risk relative to the most affluent, while they decrease their non-diversifiable risk-taking. Our evidence suggests that the least affluent victims respond to cybercrime by becoming gamblers, while the most affluent victims respond to cybercrime in a more systematic way in order to restore their pre-cybercrime wealth level.

Finally, we also shed some light on basic forensic models to gauge Ethereum addresses' prob-

ability of involvement in cybercrime. Victims can be better predicted, with a parsimonious model explaining between one-fifth and one-third of the variation in the data. Cybercriminals are harder to detect, possibly due to an effort to conceal their true intentions. Nevertheless, a parsimonious model explains roughly one-tenth to one-fifth of the variation in the data.

In what follows, we briefly relate our findings to the existing literature. Then, in Section 2, we describe our data, show aggregate statistics for cybercrime on the Ethereum blockchain, and derive our cybercrime taxonomy. Section 3 introduces our empirical design including our difference-in-differences model and matching method. Section 4 discusses our results, and Section 5 concludes.

### 1.1 Related Literature

Our study contributes to at least four different streams in the literature on cryptocurrency-related market misconduct and fraud.

First, our study contributes to the growing literature on the *financing of illegal activity through cryptocurrency*. Foley et al. (2019) document that 26% of all Bitcoin users and 46% of Bitcoin transactions are related to illegal activity on marketplaces like Silk Road where people could buy illegal drugs, pornography, and even murder-for-hire. Other studies have investigated specific forms of cybercrimes and malware. For example, Amiram et al. (2022) report that greater-than-usual blockchain activity can be linked to the vicinity of terrorist attacks, suggesting that terrorists financed the Sri Lankan Easter Bombing through cryptocurrency. Cong et al. (2022) examine, *inter alia*, dark web conversations in Russian to shed light on the organization of crypto-cybercriminals. More generally, Karapapas et al. (2020) relate the emergence of identity-concealing cryptocurrencies to the global rise of ransomware attacks. Given the heightened exposure of token investors to cybercrime, a growing literature argues that decentralized finance might benefit from more intermediaries, such as crypto funds, to manage cybercrime risk for individual investors (Cumming et al., 2022; Dombrowski et al., 2023; Fisch and Momtaz, 2020; Momtaz, 2022; Zetzsche et al., 2020).

Second, we add to numerous studies on a specific cybercrime type, namely, *pump-and-dump schemes* (e.g., Hamrick et al., 2018; Gandal et al., 2018; Li et al., 2021; Dhawan and Putniņš, 2023). The Securities and Exchange Commission (2013) and Bartoletti et al. (2020) have alerted

individuals to the presence of token-based Ponzi schemes. Market manipulations, which are deliberate and illicit actions taken by market participants to artificially alter the price of a cryptocurrency with the intention of gaining unlawful profits (Gandal et al., 2018), have been explored in several papers. For example, computer scientists have shown that smart contracts contain various vulnerabilities (e.g., Kalra et al., 2018; Luu et al., 2016; Nikolić et al., 2018), which have, for example, famously been exploited in what is known as "the DAO exploit," in which over $50 million was diverted away by fraudsters (Dhanani and Hausman, 2022). Yet other studies have taken a macro view and have categorized different fraud types associated with cryptocurrencies (Hornuf et al., 2022; Trozze et al., 2022). However, we differ from many previous studies by examining the most relevant types of fraud at the level of the individual investor address, taking into account all transactions on the Ethereum blockchain.

Third, given that our study also explores how cybercrimes influence investors' returns and risk preferences, this paper also relates to previous scholarship on returns and risk-taking of cryptocurrency investors, which has focused on the effect of market-based instruments such as derivatives (Alexander and Heck, 2020; Alexander et al., 2023; Hoang and Baur, 2020), spot trading volume and liquidity (Balcilar et al., 2017; Bouri et al., 2019; Naeem et al., 2020; Leirvik, 2022), the impact of stock, foreign exchange, and gold markets on cryptocurrency returns and volatility (Panagiotidis et al., 2018; Panagiotidis et al., 2019); and real-estate tokens (Kreppmeier et al., 2023). Liu et al. (2022) estimate long-short strategies based on ten cryptocurrency characteristics, which can be accounted for by a three-factor model that includes cryptocurrency market, size, and momentum. The impact of cybercrime on investor returns, on the other hand, has rarely been studied, much less the impact of subsequent risk-taking behavior of investors.[2] Thus, besides quantifying the extent of the fraud and developing a taxonomy of fraud on the Ethereum blockchain, our study contributes to the literature by showing how different types of scams affect the risk preferences and return prospects of investors on the Ethereum blockchain.

Finally, quantifying the extent of the fraud is important, because the effects of fraud on the Ethereum blockchain could also extend to traditional financial markets. The risk for financial institutions results from the fact that they are involved in transactions or investments on the Ethereum blockchain and may lose funds if those transactions or investments turn out to be fraudulent or

[2] A notable exception is the working paper by Fang et al., 2021.

if they become subject to cybercrimes such as ransomware attacks or hacking (Board of Governors of the Federal Reserve System et al., 2023). Cryptocurrencies that are subject to cybercrimes are also more volatile and therefore constitute more risky positions (e.g., Acemoglu et al., 2016; Caporale et al., 2020; Corbet et al., 2020; Gandal et al., 2018). If fraudulent activities on the Ethereum blockchain result in a loss of trust in the technology, then innovative projects such as blockchain-based settlements for asset trading and cryptocurrency-related business activities of financial institutions might be negatively affected. Finally, if traditional financial institutions provide services related to the Ethereum blockchain,[3] such as custody or trading, they may be exposed to risks related to the illegal activities of the blockchain's customers or counterparties. For that reason, identifying the extent and types of fraud on the Ethereum blockchain as well as the subsequent behavioral responses of investors is one of the most critical tasks to protect the reputation of this technology and to reduce the business risks for traditional and new financial organizations. After all, with increasing adoption of cryptocurrencies and smart contracts, fraud on the Ethereum blockchain can undermine the integrity of the financial system as a whole.[4]

## 2 Data and Taxonomy of Cybercrime on Ethereum

### 2.1 Data

A novelty of our empirical setting is the granularity of transaction-level data to identify victims who interacted with cybercriminals, a level of detail so far rarely examined on the Ethereum blockchain. Notable exceptions of studies using extensive on-chain data are Easley et al. (2019), Foley et al. (2019), Sokolov (2021), and Hoang and Baur (2022); however, all of these investigate transactions on the Bitcoin blockchain. Unlike Bitcoin, the Ethereum blockchain acts not only as a payment network but also as the basis for numerous decentralized applications (dApps) facilitating a more diverse range of financial activities. It is thus not surprising that the Ethereum blockchain hosts a broader range of cybercrimes. In the following section, we describe our method of identifying

[3]For example, private equity firm KKR tokenized part of its $4 billion Health Care Strategic Growth Fund II to the Avalanche blockchain, which allowed retail investors to engage in the fund. See https://www.forbes.com/sites/michaeldelcastillo/2022/09/24/kkr-blockchain-access-to-4-billion-fund-opens-door-to-crypto-investors/?sh=555c3ef84fce (retrieved July 25, 2023).

[4]For example, when First Citizens Bank recently agreed to buy most of what was left of Silicon Valley Bank, there was one thing they carved out: cryptocurrencies and loans backed by crypto. Hence, if confidence in crypto assets wanes, potentially due to fraud, the traditional financial system could also be at risk.

cybercriminals and victims, leading to our comprehensive taxonomy of cybercrime on the Ethereum blockchain.

### 2.1.1 Identifying Cybercriminals

To identify the extent of fraud on the Ethereum blockchain, we first acquired a list of blockchain addresses of cybercriminals from *Etherscan* and *Scam Alert*. *Etherscan*, a block explorer and analytics platform for Ethereum, assigns public name tags and labels to addresses that are of public interest. Any address associated with fraudulent activities has a brief warning message attached to it, providing investors with details of the purported scam. We include in our list of cybercriminals all blockchain addresses that *Etherscan* labeled as *exploit*, *hack*, *heist*, *phish*, *Ponzi scheme*, and/or *scam*. In a next step, we supplemented the list of blockchain addresses of cybercriminals with proprietary data from *Scam Alert*, which is operated by *Whale Alert*. This renowned blockchain analytics engine uncovers and tracks the activities of cybercriminals. Users can submit scam reports to and request address and website verifications from *Scam Alert*. A team of experts verifies and analyzes this information in real time.[5] The resulting list of cybercriminal blockchain addresses includes detailed information about each scam, such as the type of scam, total earnings per address, payments received, and the date of the first scam report. The list of cybercriminal blockchain addresses includes 5,644 unique addresses. Since *Etherscan* does not provide the date when a case was first reported, we manually search for the earliest relevant posts on *Twitter*, *Reddit*, and other social media platforms to identify when the incident first came to public attention.

### 2.1.2 Identifying Cybercrime Victims

As a public blockchain, Ethereum stores all transaction records on its distributed ledger, which we use as our primary data source for victim address identification. The public transaction data allows us to observe the transaction history of those who have interacted with and fallen victim to fraudsters operating on the Ethereum blockchain. Based on the identified and externally verified cybercriminal blockchain addresses, as described in Section 2.1.1, we extracted a list of addresses

[5] *Scam Alert* maintains a team of blockchain crime experts who collaborate closely with law enforcement agencies and consumer protection initiatives to detect and monitor crypto-related crime more effectively. For more information, visit here.

from all transactions on the Ethereum blockchain in which a positive sum of funds was transferred to cybercriminal addresses, starting with the Ethereum genesis block on July 30, 2015, and ending on December 31, 2021. To mitigate the risk of mistakenly identifying cybercriminals as victims, we excluded addresses repeatedly transacting with cybercriminals, since these could be potentially be involved in the cybercrime themselves. Furthermore, if the list of victim addresses appeared in the *Etherscan Public Name Tags and Labels*, we assumed that they belong to public entities and excluded them from our sample. Our final sample includes 200,865 unique victim addresses.[6]

### 2.1.3 The Evolution of Cybercrime on Ethereum

Figure 2 plots the evolution of cybercrime on the Ethereum blockchain over time. Panel A shows the dollar value of funds lost due to fraudulent transactions on the Ethereum blockchain over time. Across the entire time period analyzed, the median value of funds lost due to fraudulent activities per address is $506.76. The mean value of funds lost per address is considerably higher at $1,476.64. This skewness is driven by a number of extremely large losses by some victims. During bullish phases of the cryptocurrency market, when the value of cryptocurrencies typically increases significantly, we observe a significant rise in the mean and median values of funds lost to fraud. Panel B of Figure 2 presents the number of transactions to fraudulent accounts and the average share of blockchain address balance lost due to scam activities in the Ethereum blockchain. During phases of relative stability or downturns in the crypto market, the proportion of balance lost to scams tends to rise. In contrast, during periods of market upswings, the share of balance lost to scams appears to decrease.

**[Place Figure 2 about here.]**

[6]To effectively implement a difference-in-differences setting, victims of cybercrimes must have interacted with the fraudsters before the fraud became public so that victim behavior before and after the scam can be compared. Nonetheless, in our dataset, a non-negligible number of victims initiated transactions with fraudulent accounts even after public disclosure of the scam, a circumstance which falls outside the scope of suitability for a difference-in-differences framework, because the victim may have known the fraudster or might have even been part of the scam. Consequently, these addresses have been removed from our sample.

## 2.2 A Taxonomy of Cybercrime on the Ethereum Blockchain

Based on our sample, we derive a taxonomy of 19 unique categories of cybercrime on the Ethereum blockchain, following and extending Hornuf et al.'s (2022) approach for categorizing fraud in ICOs. We quantify the total amount of funds transferred from victims to cybercriminals' addresses and show that the 5,644 cybercriminals' addresses received an average of $1.78 million from victims' blockchain addresses, totaling $1.65 billion. This amount exceeds the self-disclosed figures of stolen funds reported by victims to the FTC by a factor of almost 16, which highlights the significant underreporting of cybercrime by regulators.

Table 1 outlines our taxonomy and describes each fraud category in detail. *Ponzi schemes*—the most common scam involving cryptocurrency, accounting for 60% of the aggregate stolen funds on Ethereum—promise high returns to investors with little or no risk. These scams are not new and constitute a digital adaptation of fraudulent activities seen in traditional finance (e.g., Hofstetter et al., 2018). The difference lies in the fact that the digital nature of blockchain technology, combined with the anonymity of blockchain addresses and lax regulatory oversight, has largely enhanced the impact and potential reach of Ponzi schemes. Almost a billion dollars were transferred into on-chain Ponzi-related accounts. Off-chain transactions and scams are also common in crypto Ponzi schemes, with the aim of soliciting funds from non-tech-savvy investors. When these off-chain Ponzi schemes are factored in, the true magnitude of the financial impact is likely even greater. One of the largest crypto Ponzi schemes, *PlusToken*, is purported to have defrauded $4 billion in 2019,[7] of which only a small part is recorded on-chain and the larger part happened off-chain in blockchain addresses of crypto exchanges.

With the rise of cryptocurrency and blockchain technology, new forms of cybercrimes have also emerged, exploiting the unique characteristics of digital assets. The *giveaway* scam is one notable example, representing the second-largest cybercrime, with 18% of the aggregate stolen funds on Ethereum. It involves the misrepresentation of the identities of reputable companies, exchanges, or influential individuals. These scams are primarily disseminated via social media platforms and are structured to mimic authentic promotions by crypto companies or exchanges.[8] In some well-known

[7] See, e.g., https://www.wsj.com/articles/cryptocurrency-scams-took-in-more-than-4-billion-in-2019-11581184800.

[8] Legitimate giveaways are often used as marketing tools to enhance brand awareness, facilitate product promotion, and drive user acquisition. They have become a popular method for engaging with potential customers while simultaneously promoting the products or services. In a legitimate giveaway, the organizer will not ask participants to send any

giveaway scams, the perpetrators imitate the largest cryptocurrency exchange, Binance, in order to inspire trust among investors.[9] Investors are then invited to send a fixed amount of cryptocurrency, usually Bitcoin or Ethereum, with the promise of high returns or rewards—a sign of an illegitimate operation. Once the investor transfers the funds, the fraudster does not uphold the initial promise. On-chain transaction data indicates that giveaway scams resulted in transfers totaling $274 million to addresses associated with this type of fraud as of the end of December 2021.

Another major scam is *exploits*, a phenomenon unique to digital systems, such as blockchains. An exploit occurs when an individual or group discovers and takes advantage of a vulnerability or bug within a system. Unlike a hack, the vulnerability is accidentally left in the code by the developer. On the blockchain, exploits often involve manipulation of smart contracts. Because they are automatically executed if certain conditions are met, a small bug or overlooked vulnerability can result in significant financial losses if exploited by malicious actors. A notable example is The DAO exploit in 2016, where an attacker exploited a code vulnerability in a decentralized autonomous organization on the Ethereum blockchain, resulting in a loss of about $50 million (Dhanani and Hausman, 2022), underscoring the potential magnitude and sophistication of blockchain exploits.

*Rug pulls* and *exit scams* are types of fraudulent activities that have specifically emerged with the advent of cryptocurrencies and ICOs. These scams involve the intentional abandonment of a project after attracting investments, often leading to significant losses for investors. In an exit scam, the founders or promoters of a blockchain project, after raising funds through an ICO, disappear with the invested capital. Rug pulls, a relatively newer form of scam, are particularly prevalent in the decentralized finance (DeFi) sector. Developers abruptly abandon a project and withdraw the liquidity from decentralized exchanges, causing a significant drop in the value of the project's token. The sudden removal of liquidity makes the tokens almost worthless, leaving investors unable to offload their holdings. In a *fake token scam* fraudsters pretend to offer well-known tokens by using similar token names and symbols. Unsuspecting users will exchange funds for worthless tokens, which have no inherent value and cannot be traded.

**[Place Table 1 about here.]**

cryptocurrency and the offerings are typically modest.

[9] See, for example https://www.binance.com/en/blog/community/know-your-scam-protect-yourself-from-binance-imposter-scams-8186206274508844717, retrieved July 31, 2023.

# 3 Empirical Design

This study aims to identify the causal impact of blockchain-related cybercrimes on victims' investment behavior. Therefore, we adopt a difference-in-differences approach that consists of a pre- versus post-cybercrime comparison between victims of cybercrime and a matched sample of non-victims/non-cybercriminals. Our main model investigates investment behavior, especially risk-taking and the risk-adjusted returns per address, based on monthly address-level panel data. We specify the following baseline regression model:

$$\begin{aligned} Y_{i,t} = \; & \beta_1 \times \mathbb{1}[\text{After cybercrime}]_{i,t} + \beta_2 \times \mathbb{1}[\text{Cybercrime victim}]_{i,t} + \\ & \beta_3 \times \mathbb{1}[\text{After cybercrime}]_{i,t} \times \mathbb{1}[\text{Cybercrime victim}]_{i,t} + \Omega_{i,t}\gamma \end{aligned} \tag{1}$$

where $i$ indexes blockchain addresses and $t$ indexes months, and $Y_{i,t}$ captures measures of risk-taking such as total risk, diversifiable risk, and non-diversifiable risk, as well as risk-adjusted returns computed as alphas, which we obtained from a cryptocurrency asset pricing model (Liu et al., 2022). $\mathbb{1}[\text{After cybercrime}]_{i,t}$ denotes an indicator that takes the value of 1 in the month of a cybercrime and thereafter, 0 otherwise. $\mathbb{1}[\text{Cybercrime victim}]_{i,t}$ denotes an indicator that takes the value of 1 if the focal blockchain address fell victim of a cybercrime and 0 if the blockchain address belongs to a matched non-victim/non-cybercriminal. $\Omega_{i,t}$ represents a matrix of controls and fixed effects, including blockchain address age, calendar-month $\times$ calendar-year fixed effects, and cybercrime-type fixed effects. Finally, note that we sample only from one-time victims in order to ensure that the identification of the average treatment effects in our model is not confounded by overlapping periods with other cybercrime events affecting the blockchain address. In our model, the average treatment effect is thus measured as the difference-in-differences estimator $\beta_3$.

## 3.1 Matching Cybercrime Victims with Non-Victims

Our study draws on the entire population of Ethereum blockchain transactions, which is high-dimensional in its scope. Due to the size and complexity of the data, there exists a notable imbalance between victims and non-victims, because the number of non-victims is substantially larger

than that of victims.[10] In the context of our study, this high dimensionality combined with the imbalance between victims and non-victims could lead to significant attenuation bias. To address this issue, we use a Euclidean distance matching procedure to match each victim's address (treatment group) to a non-victim/non-cybercriminal address (control group) that is most similar, according to a multidimensional vector consisting of blockchain address balance, blockchain address age, trading activity, and address diversification. The similarity is measured by minimizing the Euclidean distance between these vectors, based on data from the three months preceding the scam being identified to the public. This approach allows us to create a balanced representation of the treatment and control groups. Furthermore, we recognize that the substantial price fluctuations in crypto markets could affect our nominal variables; hence, we compare investor addresses that entered the market in the same month, allowing for a more balanced and fair comparison of risk and return developments.

Table 2 shows that the matching of victims and non-victims/non-cybercriminals was successful, substantially reduced the bias, and draws the sample densities of our treatment and control groups closer together. The standardized bias for each covariate is calculated as the difference in means in the treatment and control groups, divided by the standard deviation in the control group. This value is then represented as a percentage. A lower standardized % bias post-matching indicates a better balance between the treatment and control groups in terms of that specific covariate. The matching process led to significant reductions in bias for all variables. That is, our matching reduced the bias for blockchain address age by 100% (perfect matches), for blockchain address balance by 81.7% (with the remaining difference being statistically non-significant, with a p-value of 0.55), for trading activity by 97.8%, and for diversification by 96.6%.

**[Place Table 2 about here.]**

### 3.2 Matching Quality and Parallel Trends

As another plausibility check for our matching quality, we look at parallel trends in a non-matched variable. It would be reassuring if the trends between treatment and control observations are

[10] In our study, we regard each address as an individual portfolio. One limitation of this approach is that individuals can create multiple addresses; thus, one address may not fully represent an individual's portfolio or investment behavior.

parallel in a non-matched variable because this would suggest that the matching dimensions also capture more fundamental and unobserved behavior at the observational level. In particular, we plot monthly raw returns for victims and matched non-victims/non-cybercriminals in Figure 1. Barber and Odean (2000) raw returns, as defined in Table A.1, are especially well-suited to illustrate the matching quality because they do not merely reflect a single trading dimension of victims and matched non-victims/non-cybercriminals, but rather result from the cumulative investment decisions of blockchain address owners along all dimensions. Figure 1 shows reconfirming evidence that our matching was successful. Specifically, we observe parallel and mostly identical trends in the twelve months leading up to the cybercrime. However, as one would expect, in the month of the cybercrime trends start to diverge. Victims of cybercrime make investment decisions that increase their Barber and Odean (2000) raw returns relative to matched non-victims/non-cybercriminals. The cumulative effect 12 months after the cybercrime amounts to a positive return differential of 0.3% for victims of cybercrime, which is statistically highly significant.[11] That is, cybercrime impacts victims in a way that is beneficial for their raw returns.

## 3.3 Variable Construction

### 3.3.1 Outcome Variables: Risk and Risk-adjusted Return

We adopt an empirical approach based on a state-of-the-art asset pricing model to understand the risk–return dynamics at the level of individual blockchain addresses. Specifically, we estimate the cryptocurrency three-factor model developed by Liu et al. (2022) as the basis for constructing risk-related variables. This model was designed to effectively capture the unique risk factors inherent in the cryptocurrency market, which allows us to break down the *total risk* associated with an address-level portfolio into diversifiable and non-diversifiable risks at the issuer level and consequently back out *excess returns*, as measured by the alphas.

In the context of our study, we interpret these risk factors as measures for the risk preferences of victims of scams. The *total risk* of an address, defined as the overall variability of its returns, gives a sense of how much risk the investor is exposed to due to their investment choices. *Total risk* is then

[11] Note that in order not to compare apples to oranges, we consider victims' pure behavioral response and do not consider the loss caused by the scam when calculating returns. When we look at fraud losses and behavioral responses together, we find that over a 12-month period after being defrauded, cybercrime victims perform 10% worse than non-victims/non-cybercriminals.

decomposed into *diversifiable* and *non-diversifiable risk* to further understand investor behavior. *Diversifiable risk* represents the portion of risk that could be eliminated through effective portfolio diversification. High levels of diversifiable risk suggest that the investor is holding a portfolio that is not well-diversified, indicating a potential lack of risk mitigation strategies. *Non-diversifiable risk*, on the other hand, captures the inherent risk associated with the overall cryptocurrency market. It represents the systematic risk that an investor cannot reduce, regardless of how well the portfolio is diversified. High levels of non-diversifiable risk imply that the investor is taking positions in higher-risk crypto market segments.

The gross monthly portfolio *return* at the address level is computed using the beginning-of-day position statements. Following Barber and Odean (2000), we make two simplifying assumptions. First, we assume that all tokens are bought or sold at the end of the month. Second, we ignore intra-month trading. By tracking these measures over time, we can infer changes in an investor's risk appetite and evaluate how exposure to fraud events impacts portfolio returns. Table A.1 reports detailed definitions of these variables.

### 3.3.2 Other Measures of Investor Behavior

We also construct investor behavior variables at the address level using blockchain transaction data. We measure the investment horizon denoted by *churn rate*, that is, how frequently an address rotates its positions, and *diversification*, the number of unique tokens that an address holds at the end of each month. We also look at trading activity, which measures the number of transactions per address within a month. Trading activity provides insights into the investor's market engagement and potential responsiveness to fraud events.

Using the address-level monthly portfolio compositions, we also quantify the share of different classes of crypto assets, which measures the proportion of the investor's total portfolio allocated to *lottery tokens*, *stablecoins*, and *altcoins* at the end of each month. The share of different classes of crypto assets offers insights into the change in investment preferences and risk tolerance by investors with respect to their overall portfolio composition. Arguably, a larger share of Ether in an address-level portfolio is associated with more fraudulent activity, because scams are typically conducted in the native currency rather than a specific lottery or altcoin. Therefore, a higher share

of lottery tokens, stablecoins, and altcoins in an address-level portfolio is most likely associated with less fraud. All variables are defined in Table A.1 in the Appendix.

### 3.4 Summary Statistics

Table 3 presents summary statistics for blockchain addresses that fell victim to cybercrimes, reporting a broad spectrum of address characteristics. An average victimized blockchain address experienced a Barber and Odean (2000) raw return of 13.2% over the entire sample period; the median figure of 0 indicates that more than half of these addresses realized no or negative returns. Therefore, there are significant differences in investment results across different blockchain addresses.

Regarding portfolio activity, an average victimized address has a turnover rate of 5.5% and holds an average of 2.3 tokens. The size of addresses is right-skewed with an average blockchain address balance of $9,218 and a median value of $16. The age of addresses varied, averaging at 18 months, with a median age of 16 months, indicating that many victims were relatively new to the Ethereum blockchain. In terms of investment preferences, 14.6% of these addresses invested in lottery tokens, with an average portfolio share of 4.3%. Stablecoins were less popular, with just 4.8% of addresses investing in them and dedicating an average of 0.6% of their portfolio to this asset type.

Examining risk measures, the average victim blockchain address assumed a diversifiable risk of 0.311, a non-diversifiable risk of 0.095, and consequently a total risk of 0.407. Notably, despite these risks, the alpha value, which measures the blockchain address's return in excess of its expected return, averaged at a positive 6.4%. The average loading on the size factor is 0.551, which suggests that these addresses have a mild sensitivity to changes in the size factor, with a tilt towards larger cap assets. The average loading on the momentum factor is -0.370, implying that the returns of these blockchain addresses tend to move in the opposite direction to changes in the momentum factor. Thus, these blockchain addresses likely hold assets that have recently underperformed in the market.

**[Place Table 3 about here.]**

Table 4 presents the summary statistics for a matched sample of victims and non-victims/non-cybercriminals over the short-term period of three months prior to and after the scam became public. Victims showed a higher average return in both periods compared to their matched non-victims/non-cybercriminals. Before the scam became public, victims and non-victims displayed broadly similar behaviors and characteristics in several aspects. Returns, for example, were similar, with averages of 9.0% and 8.5% for victims and non-victims/non-cybercriminals, respectively. This similarity extends to metrics like trading activity and diversification, where both groups had close averages. Blockchain address balance, churn rate, and blockchain address age also were generally similar in the pre-scam phase. The average age at the time of the revelation of scams is identical, indicating that we compare addresses that entered the market in the same month, which allows us to account for the effect of macroeconomic trends on our outcome variables.

In the realm of risk factors, there was also a broad similarity between victims and non-victims/non-cybercriminals during the pre-scam phase. The means for diversifiable risk, non-diversifiable risk, total risk, market, momentum, size, and alpha were all largely similar between the treatment and control group. However, there were some minor pre-scam disparities in terms of excess returns, particularly in terms of investments in lottery tokens, stablecoins, and altcoins.

**[Place Table 4 about here.]**

Table 5 offers a comprehensive view of the mean and median summary monthly statistics for victims by the type of scams they fell victim to for the entire sample period. Regarding returns, victims of Ponzi schemes, hacking, and stolen crypto incidents reported the highest average returns at 14.5% and 14.3%, respectively, while victims of exploit and hardfork scams showed the lowest average return at 4.7%.

In terms of churn rate, victims of exploits and hardfork scams had the highest average at 0.401, while those affected by Ponzi schemes and hacks had the lowest average churn rate, indicating a lower frequency of switching from one investment to another. Diversification, a measure of the number of tokens held by an address, is highest on average for victims of fake token sales and lowest for those affected by Ponzi schemes, hacks, and stolen crypto scams. Victims of exploits and hardfork scams held the highest average balances at $176,000, while victims of hacks and stolen crypto scams had the lowest at $684. The average trading activity was highest for victims

of investment scams, while those affected by hacks and stolen crypto scams reported the lowest average trading activity.

Victims of sextortion and other scams had the oldest accounts; those who fell victim to Ponzi schemes had the youngest. Interestingly, victims of fake token sales and phishing scams were the most likely to invest in lottery tokens, with average participation shares of 73% and 67%, which constitute the highest shares of lottery token investments. Conversely, Ponzi scheme and hack victims had the lowest involvement in lottery token investments, indicating that they were generally less eager to take risky positions. Notably, individuals who fell prey to darkweb, exchange, or charity scams had a higher likelihood of having stablecoins in their portfolios and allocate a higher share of their portfolio to stablecoins. Given the nature of these scams, it is possible that these types of fraudulent activities may frequently involve or target stablecoins.

**[Place Table 5 about here.]**

# 4 Results

## 4.1 Treatment Effects of Cybercrime on Investor Risk-Taking

The graphical evidence in Figure 1 suggests that victims of cybercrime change their investment behavior in a way that increases their post-cybercrime *non*-risk-adjusted Barber and Odean (2000) monthly returns relative to matched non-victims/non-cybercriminals, not accounting for the loss that results from the scam itself. In this and the following Section 4.2, we study what drives this pure behavioral investment pattern and whether victims' post-cybercrime risk-adjusted returns are equally positive.

To this end, we estimate our main difference-in-differences model, as defined in Equation 1, with three different dependent variables: total, diversifiable, and non-diversifiable risk-taking. We also estimate the models for symmetric event windows of 3 and 12 months before and after the cybercrime event in Tables 6 and 7, respectively, accounting for the dynamic structure of how victims of cybercrime adjust their risk-taking levels. Figure 3 shows that the 3-month window is well-suited to capture the short-term response of victims to cybercrime, while the 12-month window captures a more permanent effect of cybercrime on address-level risk-taking. Finally, we estimate *average*

*treatment effects* based on all observations in our sample to quantify the aggregate impact cybercrime had on users of the Ethereum blockchain. We also estimate *heterogeneous treatment effects* for individual cybercrime categories to gauge the variance in treatment effects across different cybercrime types. All our models include granular calendar-month $\times$ calendar-year fixed effects and the model for the average treatment effect estimation also includes cybercrime-type fixed effects. Note that our analysis for the symmetric 3-month event window draws on more than 4.5 million blockchain address–month observations and the one for the symmetric 12-month event window draws on more than 7.8 million blockchain address–month observations.

The difference-in-differences results for the symmetric 3-month event window in Table 6 show the effects of cybercrime on blockchain address-level total, diversifiable, and non-diversifiable risk-taking in Panels A, B, and C, respectively. Again, the first two models estimate *average treatment effects* and models (3) to (11) estimate *heterogeneous treatment effects*.

Comparing the short-term effects for the 3-month event window with the more long-term effects for the 12-month event window is interesting because many identified effects are reversed. For example, the average treatment effects in the between-cybercrime model in column (1) are negative (–0.0157), positive (0.0043), and negative (–0.0200) for the 3-month window in Table 6 and positive (0.0230), negative (–0.0027), and positive (0.0257) for the 12-month window in Table 7 for total (Panels A), diversifiable (Panels B), and non-diversifiable risk-taking (Panels C), respectively. The structure of the average treatment effects is similarly reversed in the within-cybercrime model in column (2). For brevity, we only provide an overarching comparison of the heterogeneous treatment effects. A number of cybercrimes entail similar effects over the 3- and 12-month event windows, although the longer window mostly exhibits stronger effects in terms of both statistical and economic magnitude. In particular, the effects for Ponzi schemes, hacks, and exploits are consistent across the event windows. In contrast, several cybercrimes yield either reversed effects or are non-significant in the shorter event window. These include giveaways, phishing scams, investment scams, fake token scams, darkweb shop-related cybercrime, and sextortion.

**[Place Table 6 about here.]**

Table 7 presents the difference-in-differences results for the symmetric 12-month event window for blockchain address-level total, non-diversifiable, and diversifiable risk-taking in Panels A, B,

and C, respectively. Our first two models estimate *average treatment effects*. In Panel A (total risk), the *between*-cybercrime-type model (i.e., the model without cybercrime-type fixed effects) in the first column estimates an average treatment effect of 0.023, which is statistically highly significant at the 0.1% level. In economic terms, the estimate suggests that, given the full-samples average total risk-taking of 0.407, victims of any type of cybercrime increase their total risk-taking levels by 5.7% (= 0.023 / 0.407) as a response to the scam event. The average treatment effect in the *within*-cybercrime-type model (i.e., the model with cybercrime-type fixed effects) in the second column is only 0.0033; that is, victims increase post-cybercrime risk-taking level by 8.1%, albeit the within-cybercrime-type effect is statistically non-significant. The non-significant effect on the total risk-taking level is caused by counteracting effects for non-diversifiable (positive effects) and diversifiable (negative effects) risk-taking levels in Panels B and C, respectively. Panel B (non-diversifiable risk) shows that post-cybercrime victims reduce their blockchain address-level non-diversifiable risk-taking. The highly statistically significant difference-in-differences estimators of –0.0027 and –0.0142 for the between- and within-cybercrime-type models suggest that victims decrease their non-diversifiable risk-taking by 0.8% and 4.6%, respectively, given the full-sample non-diversifiable risk-taking average of 0.311. Panel C (diversifiable risk) shows that post-cybercrime victims increase their blockchain address-level diversifiable risk-taking in the long term. The highly statistically significant difference-in-differences estimates of 0.0257 and 0.0175 for the between- and within-cybercrime-type models suggest that victims increase their diversifiable risk-taking by 27.1% and 18.4%, respectively, given the full-sample diversifiable risk-taking average of 0.095.

The models in columns (3) to (11) contain *heterogeneous treatment effects* by cybercrime type. Overall, we find that post-cybercrime blockchain address-level risk-taking critically depends on the type of cybercrime a victim fell for.

Cybercrime that leads to an increase in victims' total risk-taking levels are Ponzi schemes, events on the darkweb, and sextortion. Ponzi scheme-related cybercrime increases victims' total risk-taking levels by 14.8% (= 0.0604 / 0.407), which is associated with a reduction in non-diversifiable risk-taking and an increase in diversifiable risk-taking of –23.1% (= –0.0219 / 0.095) and 26.4% (= 0.0822 / 0.311), respectively. Darkweb-related cybercrime increases victims' total risk-taking levels by 114.6% (= 0.4665 / 0.407), which is associated with increases in non-diversifiable risk-taking and diversifiable risk-taking of 156.9% (= 0.1519 / 0.095) and 101.2% (= 0.3147 / 0.311),

respectively. Sextortion-related cybercrime increases victims' total risk-taking levels by 114.6% (= 0.4665 / 0.407), which is associated with increases in non-diversifiable risk-taking and diversifiable risk-taking of 64.9% (= 0.0617 / 0.095) and 21.0% (= 0.0653 / 0.311), respectively.

Cybercrime does not alter victims' total risk-taking levels if the event was a phishing scam or involved a fake token. At least for phishing scams, the non-significant total risk-taking effect is statistically non-significant, while the non-diversifiable and diversifiable effects are statistically significant. Victims of phishing scams increase their non-diversifiable risk-taking level by 162.1% (= 0.0154 / 0.095), while they reduce their diversifiable risk-taking level by 212.9% (= 0.0662 / 0.311).

Cybercrime that leads to a reduction in victims' total risk-taking levels are giveaways, investment scams, hacks, and exploits. Giveaway-related cybercrime reduces victims' total risk-taking levels by –12.2% (= –0.0495 / 0.407), which is associated with an increase in non-diversifiable risk-taking and a reduction in diversifiable risk-taking of 9.4% (= 0.0089 / 0.095) and –21.3% (= –0.0662 / 0.311), respectively. Investment scam-related cybercrime reduces victims' total risk-taking levels by –58.7% (= –0.239 / 0.407), which is associated with an increase in non-diversifiable risk-taking and a reduction in diversifiable risk-taking of 29.1% (= 0.0276 / 0.095) and –85.7% (= –0.2666 / 0.311), respectively. Hack-related cybercrime reduces victims' total risk-taking levels by –52.8% (= –0.215 / 0.407), which is associated with reductions in non-diversifiable risk-taking and diversifiable risk-taking of –13.7% (= –0.013 / 0.095) and –220.1% (= –0.6846 / 0.311), respectively. Exploit-related cybercrime reduces victims' total risk-taking levels by –167.5% (= –0.6818 / 0.407), which is associated with an increase in non-diversifiable risk-taking and a reduction in diversifiable risk-taking of 2.9% (= 0.0028 / 0.095) and 220.1% (= –0.6846 / 0.311), respectively.

**[Place Table 7 about here.]**

The results from Tables 6 and 7 suggest that post-cybercrime risk-taking changes with time. This result is consistent with recent literature showing that investor behavior changes over time and that, for example, the perceived risk of fraud can itself be a determinant of risk-taking by investors (Dicle, 2019). To shed light on the dynamics of the treatment effects, we plot the difference-in-difference estimates for total risk, non-diversifiable risk, and diversifiable risk in Panels A, B, and C of Figure 3. Total and diversifiable risk follow similar trends. Cybercrime reduces risk-taking along

these two dimensions for the first 5 months after the fraud became public, and, starting in month 6, risk-taking starts to climb back to the pre-cybercrime level, which it reaches around months 10 to 12. Ultimately risk-taking exceeds the pre-fraud level, which has been normalized in Figure 3 by construction to 0%, and then permanently stays at a constantly higher level after months 12 to 15. For non-diversifiable risk, the pattern is significantly different. Post-cybercrime non-diversifiable risk-taking decreases constantly over the first 12 months post-event, after which it starts to recover slowly, though never returning to the pre-cybercrime level over the 24-month observation period.

**[Place Figure 3 about here.]**

## 4.2 Treatment Effects of Cybercrime on Risk-Adjusted Returns

The evidence in the preceding section indicates that cybercrime victims, on average, reduce risk-taking relative to the pre-event level in the first year following the event. A natural next question is whether and how the adjustment to risk-taking levels is reflected in victims' risk-adjusted returns. Panels A and B of Table 8 present the results for the 3- and 12-month event windows, respectively.

Columns (1) and (2) of Table 8 report the *average treatment effects* from the difference-in-differences analyses for blockchain address-level alphas. Note that the results are consistent throughout all four model specifications, regardless of whether we look at between- or within-cybercrime-type models, or the different event windows. The coefficients are all highly statistically significant and range from –0.0353 (within-cybercrime-type model; 3-month window) to –0.0617 (between-cybercrime-type model; 12-month window). In economic terms, victims' risk-adjusted returns in the post-cybercrime period reduce by 55.2 percentage points (= –0.0353 / 0.064) to 96.4 percentage points (= –0.0617 / 0.064) relative to matched non-victims/non-cybercriminals.

The *heterogeneous treatment effects analyses by cybercrime type* again suggest that victims' risk-adjusted returns can be both positively and negatively impacted by the various categories of cybercrime. For the symmetric 12-month event window, cybercrime categories that have a *positive* effect on victims' post-event risk-adjusted returns are fake token scams, darkweb activity, and sextortion. Fake token-related cybercrime increases victims' risk-adjusted returns in the post-event period by 26.4 percentage points (= 0.0169 / 0.064). Darkweb-related cybercrime increases victims' risk-adjusted returns in the post-event period by 45.9 percentage points (= 0.0294 / 0.064).

Sextortion-related cybercrime increases victims' risk-adjusted returns in the post-event period by 44.7 percentage points (= 0.0286 / 0.064).

For the symmetric 12-month event window, cybercrime categories that have a *negative* effect on victims' post-event risk-adjusted returns are Ponzi schemes, phishing scams, investment scams, hacks, and exploits. Ponzi scheme-related cybercrime increases victims' risk-adjusted returns in the post-event period by 124.5 percentage points (= –0.0797 / 0.064). Phishing scam-related cybercrime increases victims' risk-adjusted returns in the post-event period by 29.1 percentage points (= –0.0186 / 0.064). Investment scam-related cybercrime increases victims' risk-adjusted returns in the post-event period by 18.8 percentage points (= –0.012 / 0.064). Hack-related cybercrime increases victims' risk-adjusted returns in the post-event period by 36.7 percentage points (= –0.0235 / 0.064). Exploit-related cybercrime increases victims' risk-adjusted returns in the post-event period by 298.6 percentage points (= –0.1911 / 0.064).

Unlike for victims' post-cybercrime risk-taking levels, the treatment effects on victims' risk-adjusted returns is relatively consistent throughout the symmetric 3- and 12-month event windows. Therefore, we only briefly discuss commonalities and differences at an overarching level, without going into detail. In general, the 3-month model yields slightly smaller treatment effects than the 12-month model. This can be explained by the time-series pattern in Figure 4 below, which suggests that cybercrime impacts victims' alphas permanently negatively, with the treatment effect reaching its peak around month 10 after the cybercrime. A few heterogeneous treatment effects differ for the two event windows. While investment scams and fake token scams significantly reduce victims' alphas over the 12-month window, the treatment effects are statistically non-significant for the 3-month event window.

**[Place Table 8 about here.]**

Figure 4 plots the average treatment effects from the difference-in-differences model for the monthly alphas for the 1- to 24-month post-cybercrime period. That is, we estimate our main model with alphas as the dependent variable for 24 different event windows. Figure 4 illustrates that victims' post-cybercrime alphas take a strong hit of around –3% in the month right after the cybercrime and then continue to decline to slightly more than –6% in month 10, and thereafter remain relatively stable at that level.

**[Place Figure 4 about here.]**

### 4.3 Investor Behavior, Risk-Taking, and Returns

Given that the collective evidence so far suggests that cybercrime victims increase total risk-taking and have lower risk-adjusted returns, a natural next question to investigate is which dimensions of investor behavior help explain these patterns. To this end, we simultaneously regress several characteristics of investor behavior on risk-taking and risk-adjusted returns in a correlational triple differences model. Specifically, we are interested in the triple interactions of the victim and post-scam indicators with the measures of risk-taking and risk-adjusted returns. To explain the diverging patterns for risk-taking and risk-adjusted returns in the period following a cybercrime, we would expect that at least some investor behavior characteristics load positively for the risk-related triple difference estimates and negatively for the return-related triple difference estimates, and vice versa. As measures of investor behavior, we explore blockchain address-level trading activity, churn rate, diversification, lottery token, stablecoin, and altcoin blockchain address weights, which we define in Section 3.3. Table 9 presents the regression results for the 3- and 12-month event windows in Panels A and B, respectively.

Overall, our results yield three important insights. First, *trading-related investor behavior*—that is, trading frequency and investment horizon—appears to be the primary driver of the negative treatment effects on alphas for victims in the post-cybercrime period. Second, *investment strategy-related investor behavior*—that is, diversification and ownership of different token categories—appears to be the main factor behind the increases in alphas and non-diversifiable risk and the reduction in diversifiable risk. Third, alpha and risk-taking are explained by these investor behavior measures with heterogeneous quality. In terms of the adjusted R-squared (for the 12-month window in Panel B), the altcoin weight (adjusted R-squared of 29.3%) and diversification (adjusted R-squared of 13.3%) are most meaningful in terms of the variation explained by alpha and risk-taking in these variables, followed by the lottery token blockchain address weight (adjusted R-squared of 8.0%), churn rate (adjusted R-squared of 5.3%), the stablecoin blockchain address weight (adjusted R-squared of 2.7%), and trading activity (adjusted R-squared of 0.6%).

More precisely, Table 9 shows that (i) trading activity loads significantly negatively on alpha

but has no significant relation with diversifiable or non-diversifiable risk-taking, while (ii) churn rate, our measure of how quickly investors rotate their portfolio, loads significantly negatively on alpha and diversifiable risk-taking and significantly positively on non-diversifiable risk-taking. Both results suggest that their risk-adjusted returns are falling as investors trade more after being hit by a scam. This finding is consistent with findings from Odean and Barber (1999) for traditional capital markets that investors who trade more tend to underperform. Moreover, (iii) diversification, (iv) stablecoin blockchain address weight, (v) altcoins blockchain address weight, and (vi) lottery token blockchain address weight all load significantly positively on alpha and non-diversifiable risk-taking but significantly negatively on diversifiable risk-taking. The fact that diversification, which can also be represented by a higher share of lottery tokens, stablecoins, and altcoins, increases returns and negatively relates to diversifiable risk-taking appears intuitive. The fact that lottery tokens, stablecoins, and altcoins often represent early investments, especially when compared to the native cryptocurrency Ether, and in many cases offer unique DeFi use cases, could explain excess returns represented by larger alphas. However, the fact that a higher proportion of lottery tokens, stablecoins, and altcoins are positively associated with undiversifiable risk-taking may be due to the fact that these tokens are associated with other forms of undiversifiable risk. These forms of market-wide risk may stem from investors' doubts about the ability of stablecoin issuers to maintain a currency peg, which consequently raises doubts about the premise of stablecoins and blockchain technology in general.[12]

Two observations support the meaningfulness of our results. First, the positive relation between alpha and non-diversifiable risk-taking and the negative relation between alpha and diversifiable risk-taking are consistent with arbitrage pricing theory (Fama and French, 1992, 1993; Roll and Ross, 1980). Second, although we document a dynamic structure of risk-taking levels on the event window in the post-cybercrime period, the coefficients in Panel A (3-month window) and Panel B (12-month window) in Table 9 are largely consistent, suggesting that these investor behaviors drive alphas and risk-taking independent of the observation period, reflecting some fundamental associations.

**[Place Table 9 about here.]**

[12]For a comprehensive list of failed stablecoins, see https://chainsec.io/failed-stablecoins/

### 4.4 Heterogeneous Treatment Effects by Blockchain Address Balance

Do affluent blockchain addresses react differently to cybercrime than non-affluent ones? To address the question, Figure 5 plots treatment effects for alphas, total risk, diversifiable risk-taking, and non-diversifiable risk-taking for the top 10% richest and bottom 10% poorest blockchain addresses as measured by blockchain address balance in the month prior to the focal cybercrime. Poor blockchain addresses yield significantly lower alphas over the 24 months following the cybercrime. For example, two years after the cybercrime, blockchain addresses of the richest victims yield an alpha of –4% relative to non-victim/non-cybercriminal matched control blockchain addresses, while blockchain addresses of the poorest victims yield an alpha of –5.5% relative to non-victim/non-cyber-criminal matched control blockchain addresses.

The heterogeneous treatment effect on risk-adjusted returns between rich and poor cybercrime victims can be explained by different levels of risk-taking in the post-scam period. Rich blockchain addresses take substantially less total risk than poor blockchain addresses, although risk loading by rich vs. poor blockchain addresses differs by risk type. That is, rich blockchain addresses take on less diversifiable risk and more non-diversifiable risk than poor blockchain addresses as a response to a cybercrime event. Hence, financially vulnerable investors have historically faced a dual burden: first, losing their funds to scams, and second, compounding their losses further by adopting a speculative approach after falling victim to the scam. Consequently, the role of regulators and consumer authorities becomes doubly crucial in combating cryptocurrency scams and protecting these vulnerable individuals.

**[Place Figure 5 about here.]**

### 4.5 Predicting Different Participants on the Ethereum Blockchain

#### 4.5.1 Predicting Cybercriminals

Table 10 Panel A shows blockchain address characteristics predicting cybercriminals across various crime categories. First, the age of the blockchain address does not appear to have predictive power of cybercriminals, while addresses engaged in these illicit activities tend to diversify their assets. It is technically feasible to transfer ownership of an existing blockchain address, and darkweb markets

may have emerged allowing cybercriminals to impersonate an old blockchain address if they do not already have one. The diversification might be a direct consequence of the nature of the crimes committed. If a criminal engages in various types of fraud that yield different types of tokens, this will naturally lead to a more diversified portfolio relative to those who do not. This theory could hold particularly true for cases where cybercriminals accept or demand payment in the victims' tokens, leading to an assortment of different assets in their portfolios. Moreover, stolen tokens and hacks, for example, could naturally lead to greater diversification in cybercriminals' portfolios, while cybercriminals that distribute malware often only accept a few cryptocurrencies.

Lottery token and stablecoin share are both negatively associated with criminals' block-chain addresses in all types of cybercrimes. In other words, cybercriminals appear to be highly discriminating in their choice of cryptocurrencies, avoiding both extremes of the spectrum (i.e., highly regulated stablecoins and overly speculative assets such as lottery tokens). While stablecoins provide a certain level of predictability due to their regulation and stability, their enhanced traceability and centralization may deter cybercriminals who value anonymity and control. On the other hand, lottery tokens, often associated with high-risk, high-reward speculative investing, could pose a significant risk even for cybercriminals. Despite the potential for high returns, the extreme volatility and uncertain nature of such assets could lead to substantial losses. Furthermore, the relative lack of establishment and recognition of these tokens might pose challenges in terms of liquidity and ease of transaction, making them less suitable for illicit activities.

In contrast to stablecoins and lottery tokens, we find a positive relationship between the share of altcoins in the criminals' blockchain addresses for most cybercrime types. Altcoins may offer a balance between anonymity, risk, and reward that may be appealing to cybercriminals. Unlike stablecoins, they are typically not as heavily regulated. They are also more established and less speculative than lottery tokens, reducing the risk of substantial losses due to volatility. Importantly, the nature of altcoins may provide opportunities for exploitation by cybercriminals. For example, many startups in the blockchain space often raise funds through ICOs or similar mechanisms, where they sell tokens to early investors. These tokens can sometimes be obtained in significant volumes and at lower prices during these initial phases, making them attractive to cybercriminals. Moreover, although these tokens are not as widely accepted as more established cryptocurrencies, they often have sufficient liquidity for criminals to convert them into other assets or fiat currency when

needed.

**[Place Table 10 about here.]**

### 4.5.2 Predicting Victims

Table 10 Panel B shows blockchain address characteristics predicting victims across various scam categories. The age of the blockchain address is negatively associated with victimization across all types of cybercrimes, which suggests that older and potentially more experienced investors are less likely to fall victim to cybercrimes. This could be due to increased knowledge and understanding of the cryptocurrency ecosystem and the respective risks that occur over time.

Victim addresses also tend to hold a more diversified portfolio across all cybercrime types. A potential explanation could be that individuals with more diverse portfolios might be more active in the cryptocurrency market, engaging in more transactions and with a wider array of tokens, thereby increasing their exposure to scams and other forms of cybercrime. Similar to the pattern observed among cybercriminals, the victims' share of lottery tokens and stablecoins is negatively related to all types of cybercrimes. This could suggest that victims, like criminals, also tend to avoid highly volatile assets like lottery tokens, perhaps due to their speculative nature and high risk. Additionally, victims typically hold a smaller share of stablecoins in their blockchain addresses. This could be due to their primary engagement with cryptocurrencies not being DeFi activities or trading, which often employ stablecoins. This difference in usage could affect their interaction with the blockchain ecosystem and thus their likelihood of becoming victims of various types of cybercrimes.

Interestingly, there is a positive relationship between victims' share of altcoins and most types of cybercrime. This is likely due to the relative novelty and potential for high returns from these assets. Victims may be lured by the prospect of quick profits from newly established cryptocurrencies, making them more vulnerable to scams and other forms of cybercrime.

# 5 Conclusion

This article is among the first to provide a comprehensive analysis of cybercrime on the Ethereum blockchain. We identify more than 1.78 million transactions that are externally verified to be linked to cybercrime, corresponding to an aggregate amount of $1.65 billion of funds lost. In a first step, our analysis shows that the FTC understates the amount of abducted funds on the Ethereum blockchain by a staggering factor of 16. Furthermore, our data enables us to develop a taxonomy grounded in the economic impact of each cybercrime, yielding 19 overarching categories. With the data and taxonomy in hand, we develop a causal approach to estimating how cybercrime impacts victims' risk-taking, risk-adjusted returns, and investor behavior. Using a difference-in-differences approach on victim and non-victim/non-cybercriminal matched addresses, we find that victims increase their overall risk-taking, which leads to higher Barber and Odean (2000) raw returns and lower risk-adjusted returns as measured by alphas from a state-of-the-art crypto factor model.

We find heterogeneous post-cybercrime risk-taking effects. Although total risk increases, a risk decomposition leads to higher diversifiable risk-taking and lower non-diversifiable risk-taking at the address level in the long term. We also evidence time-dependencies: diversifiable risk-taking decreases in the short term and increases permanently in the long term, while non-diversifiable risk-taking decreases in the short and medium term, but does not return to pre-cybercrime levels within a 24-month period. We show that various measures for investor behavior, including trading behavior and investment strategy, explain the differential impact of cybercrime on risk-taking and risk-adjusted returns. Finally, in post-hoc additional analysis, we show that victim and cybercrime addresses differ systematically, leading to variation that can be exploited in predictive models to screen for cybercriminals ex ante.

## References


Acemoglu, D., Malekian, A., & Ozdaglar, A. (2016). Network security and contagion. *Journal of Economic Theory*, *166*, 536–585.

Alexander, C., Deng, J., Feng, J., & Wan, H. (2023). Net buying pressure and the information in bitcoin option trades. *Journal of Financial Markets*, *63*, 100764.

Alexander, C., & Heck, D. F. (2020). Price discovery in bitcoin: The impact of unregulated markets. *Journal of Financial Stability*, *50*, 100776.

Amiram, D., Jørgensen, B. N., & Rabetti, D. (2022). Coins for bombs: The predictive ability of on-chain transfers for terrorist attacks. *Journal of Accounting Research*, *60*(2), 427–466.

Balcilar, M., Bouri, E., Gupta, R., & Roubaud, D. (2017). Can volume predict bitcoin returns and volatility? a quantiles-based approach. *Economic Modelling*, *64*, 74–81.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, *55*(2), 773–806.

Bartoletti, M., Carta, S., Cimoli, T., & Saia, R. (2020). Dissecting ponzi schemes on ethereum: Identification, analysis, and impact. *Future Generation Computer Systems*, *102*, 259–277.

Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, & Office of the Comptroller of the Currency. (2023). Joint statement on crypto-asset risks to banking organizations. *Available at https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20230103a1.pdf*.

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.

Caporale, G. M., Kang, W.-Y., Spagnolo, F., & Spagnolo, N. (2020). Non-linearities, cyber attacks and cryptocurrencies. *Finance Research Letters*, *32*, 101297.

Cong, L. W., Harvey, C. R., Rabetti, D., & Wu, Z.-Y. (2022). An anatomy of crypto-enabled cybercrimes. *Available at SSRN 4188661*.

Corbet, S., Cumming, D. J., Lucey, B. M., Peat, M., & Vigne, S. A. (2020). The destabilising effects of cryptocurrency cybercriminality. *Economics Letters*, *191*, 108741.

Cumming, D., Hornuf, L., Karami, M., & Schweizer, D. (2021). Disentangling crowdfunding from fraudfunding. *Journal of Business Ethics*, 1–26.

Cumming, D. J., Dombrowski, N., Drobetz, W., & Momtaz, P. P. (2022). Decentralized finance, crypto funds, and value creation in tokenized firms. *Crypto Funds, and Value Creation in Tokenized Firms (May 7, 2022)*.

Dhanani, A., & Hausman, B. J. (2022). Decentralized autonomous organizations. *Intellectual Property & Technology Law Journal*, *34*, 3–9.

Dhawan, A., & Putniņš, T. J. (2023). A new wolf in town? pump-and-dump manipulation in cryptocurrency markets. *Review of Finance*, *27*(3), 935–975.

Dicle, M. F. (2019). Increasing return response to changes in risk. *Review of Financial Economics*, *37*(1), 197–215.

Dombrowski, N., Drobetz, W., & Momtaz, P. P. (2023). Performance measurement of crypto funds. *Economics Letters*, *228*, 111118.

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

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fang, Y., Chen, C. Y.-H., & Jiang, C. (2021). A fight-to-safety from bitcoin to stock markets: Evidence from cyber attacks. *Available at SSRN 3864561*.

Fisch, C., & Momtaz, P. P. (2020). Institutional investors and post-ico performance: An empirical analysis of investor returns in initial coin offerings (icos). *Journal of Corporate Finance*, *64*, 101679.

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

Gandal, N., Hamrick, J., Moore, T., & Oberman, T. (2018). Price manipulation in the bitcoin ecosystem. *Journal of Monetary Economics*, *95*, 86–96.

Gaspar, J.-M., Massa, M., & Matos, P. (2005). Shareholder investment horizons and the market for corporate control. *Journal of Financial Economics*, *76*(1), 135–165.

Guiso, L., & Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, *6*(6), 1109–1150.

Hamrick, J., Rouhi, F., Mukherjee, A., Feder, A., Gandal, N., Moore, T., & Vasek, M. (2018). The economics of cryptocurrency pump and dump schemes. *Available at SSRN 3310307*.

Hoang, L. T., & Baur, D. G. (2020). Forecasting bitcoin volatility: Evidence from the options market. *Journal of Futures Markets*, *40*(10), 1584–1602.

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

Hofstetter, M., Mejıa, D., Rosas, J. N., & Urrutia, M. (2018). Ponzi schemes and the financial sector: Dmg and drfe in colombia. *Journal of Banking & Finance*, *96*, 18–33.

Hornuf, L., Kück, T., & Schwienbacher, A. (2022). Initial coin offerings, information disclosure, and fraud. *Small Business Economics*, *58*(4), 1741–1759.

Kalra, S., Goel, S., Dhawan, M., & Sharma, S. (2018). Zeus: Analyzing safety of smart contracts. *Ndss*, 1–12.

Karapapas, C., Pittaras, I., Fotiou, N., & Polyzos, G. C. (2020). Ransomware as a service using smart contracts and ipfs. *2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 1–5.

Kreppmeier, J., Laschinger, R., Steininger, B. I., & Dorfleitner, G. (2023). Real estate security token offerings and the secondary market: Driven by crypto hype or fundamentals? *Journal of Banking & Finance*, 106940.

Leirvik, T. (2022). Cryptocurrency returns and the volatility of liquidity. *Finance Research Letters*, *44*, 102031.

Li, T., Shin, D., & Wang, B. (2021). Cryptocurrency pump-and-dump schemes. *Available at SSRN 3267041*.

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

Luu, L., Chu, D.-H., Olickel, H., Saxena, P., & Hobor, A. (2016). Making smart contracts smarter. *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, 254–269.

Momtaz, P. P. (2022). Is decentralized finance (defi) efficient? *Available at SSRN 4095397*.

Naeem, M., Bouri, E., Boako, G., & Roubaud, D. (2020). Tail dependence in the return-volume of leading cryptocurrencies. *Finance Research Letters*, *36*, 101326.

Nikolić, I., Kolluri, A., Sergey, I., Saxena, P., & Hobor, A. (2018). Finding the greedy, prodigal, and suicidal contracts at scale. *Proceedings of the 34th annual computer security applications conference*, 653–663.

Odean, T., & Barber, B. (1999). The courage of misguided convictions: The trading behavior of individual investors. *Financial Analyst Journal*, 41–55.

Panagiotidis, T., Stengos, T., & Vravosinos, O. (2018). On the determinants of bitcoin returns: A lasso approach. *Finance Research Letters*, *27*, 235–240.

Panagiotidis, T., Stengos, T., & Vravosinos, O. (2019). The effects of markets, uncertainty and search intensity on bitcoin returns. *International Review of Financial Analysis*, *63*, 220–242.

Roll, R., & Ross, S. A. (1980). An empirical investigation of the arbitrage pricing theory. *The journal of finance*, *35*(5), 1073–1103.

Securities and Exchange Commission. (2013). Ponzi schemes using virtual currencies. *SEC Pub. No. 153 (7/13)*.

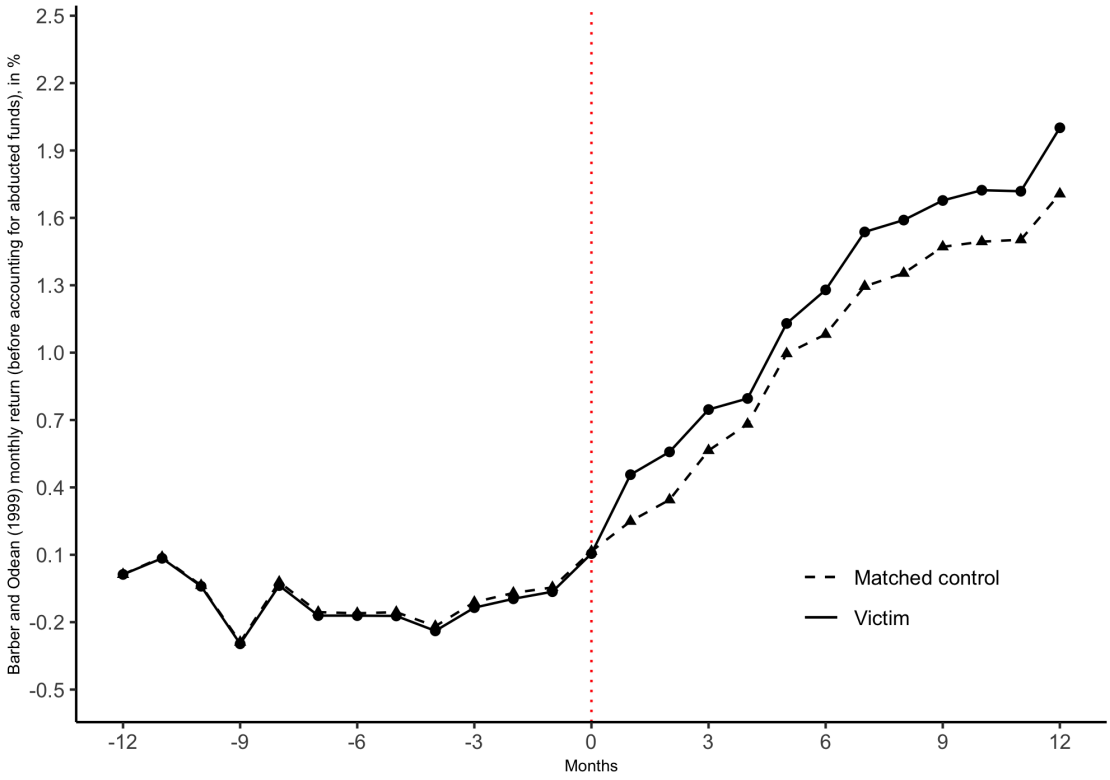
Sokolov, K. (2021). Ransomware activity and blockchain congestion. *Journal of Financial Economics*, *141*(2), 771–782.

Trozze, A., Kamps, J., Akartuna, E. A., Hetzel, F. J., Kleinberg, B., Davies, T., & Johnson, S. D. (2022). Cryptocurrencies and future financial crime. *Crime Science*, *11*, 1–35.

Zetzsche, D. A., Arner, D. W., & Buckley, R. P. (2020). Decentralized finance (defi). *Journal of Financial Regulation*, *6*, 172–203.

## Exhibits

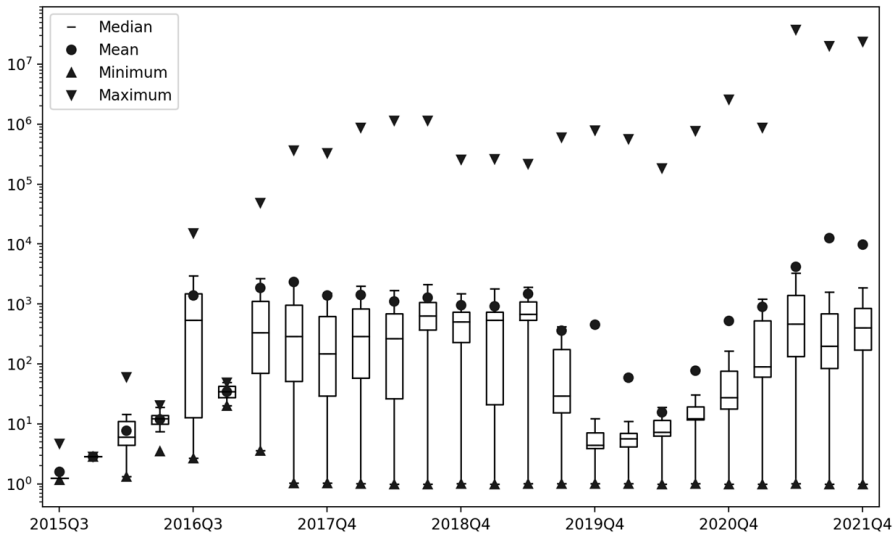
Figure 1: **Parallel trends and treatment effect of cybercrime on victims' *raw* returns**



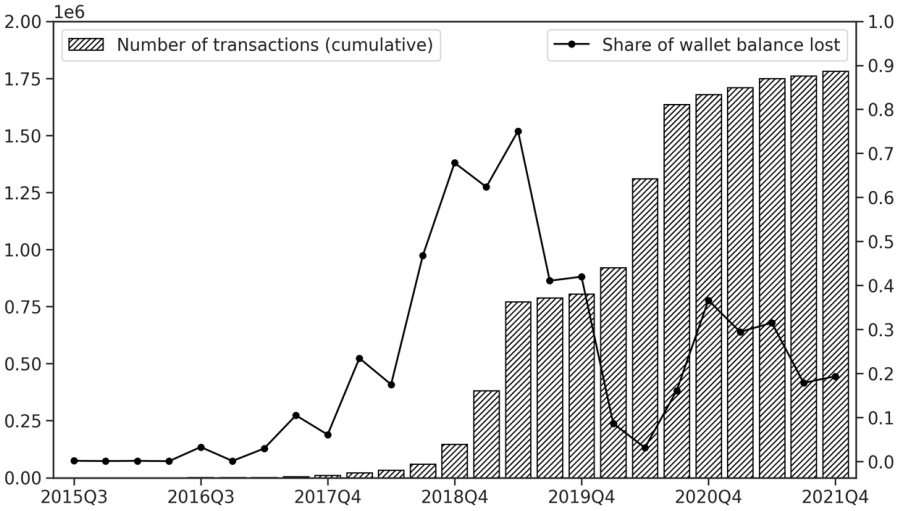

**Note:** This graph plots Barber and Odean (2000) monthly raw returns, as defined in Table A.1, for victims and matched non-victims for the time period of –12 to +12 months with respect to the focal cybercrime. Barber and Odean (2000) monthly raw returns only account for the *behavioral* effect of cybercrime on returns, not for the abducted funds due to the cybercrime *per se*. The graph illustrates that victims become better investors post-cybercrime. However, if one were to account for the nominal value of abducted funds due to the cybercrime, cybercrime victims have lost 10% of their wealth 12 months after the cybercrime relative to matched non-victims. Thus, on average, cybercrime victims lose one-tenth of their address-level wealth in a cybercrime. Note that the plot illustrates nearly perfect parallel trends for the treatment (cybercrime victims) and control observations (matched non-victims), suggesting the identification of a causal effect of cybercrime on victim behavior.

Figure 2: **Cybercrime on the Ethereum blockchain**

**Panel A:** Funds transferred to fraudulent accounts per blockchain address, in $




**Panel B:** Transactions to fraudulent addresses and the share of funds lost due to scams




**Note:** The first figure shows the interquartile range, mean, and the maximum and minimum amounts of funds transferred to fraudulent accounts. The second figure shows the cumulative sum of the number of transactions to fraudulent accounts (left axis) and the share of blockchain address balance lost due to scams (right axis).

Table 1: **A taxonomy of cybercrime on Ethereum**

| Scam category | Description | # addresses | # transactions | $ received |
|---|---|---|---|---|
| Ponzi Scheme | A type of investment fraud whereby cybercriminals lure investors with purportedly high returns with little to no risk. Without real underlying businesses, it focuses mainly on attracting new investors to make promised payments to existing investors. | 124 | 1,539,927 | 989,408,032 |
| Giveaway | A scammer poses as a major company, exchange, or celebrity hosting a giveaway and promises to send back, for instance, double the amount received from the investor. As one of the most prevalent forms of scam, it is often advertised on social media platforms. | 1,914 | 18,814 | 297,119,907 |
| Exploit | Instances where an exploiter takes advantage of a vulnerability or bug to cause unintended or unanticipated behavior to occur. | 51 | 3,099 | 214,021,559 |
| General Phishing Scam | A type of social engineering where a fraudulent message is sent by an attacker in an attempt to gain access to private data, resulting in stolen funds. When a scam lacks information on a specific fraud type, it is included in our data set as a general phishing scam. | 2,453 | 93,558 | 80,618,544 |
| Hack | An attempt to gain access to private data, which can range from stolen private keys to illegitimate or counterfeit hardware wallets, designed to steal funds. | 110 | 87,247 | 22,269,413 |
| Exchange | Fake cryptocurrency exchanges posing as legitimate exchanges. Trading volumes on these exchanges are often manipulated to appear credible. Users may be lured with additional giveaway tokens. Once the money is received by scam exchanges, users are in many cases burdened with high fees and/or denied crypto withdrawals. | 113 | 10,529 | 11,092,274 |
| Stolen crypto | Instances whereby users had their private key stolen, or their wallets hacked. | 279 | 9,893 | 9,917,668 |
| Investment | Cybercriminals pose as investment managers and contact victims offering crypto investment products. They often require an upfront fee and may also ask for private information to get access to the user's assets. | 313 | 12,515 | 9,773,136 |
| Rug Pull/Exit Scam/ICO Scam | Exit scammers are protocol founders or promoters who, during or after an initial coin offering (ICO), disappear with funds raised by investors. A rug pull is a newer form of exit scam where developers abandon a project and pull liquidity away from decentralized exchanges entirely, causing the token value to plummet to zero. | 39 | 798 | 5,671,271 |

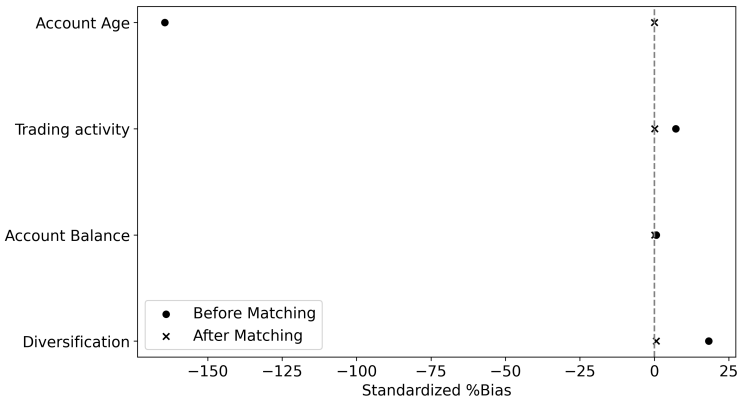
| | | | | |
|---|---|---|---|---|
| Fake Token Scam | Tokens that pose as well-known tokens by using similar token names and symbols. Unsuspecting users will exchange them using real tokens. These scam tokens usually have no value and cannot be traded. | 35 | 1,280 | 2,974,719 |
| Malware | A type of phishing scam where malicious software is planted into a device in order to gain access to the user's funds. | 18 | 1,320 | 1,871,751 |
| Fake Token Sale Scam | Scams propagated through malicious advertisements that imitate legitimate new token launches. Scammers may also pose as well-known entities, promoting fake new token sales tricking investors into purchasing their new fraudulent crypto tokens. | 21 | 373 | 956,928 |
| Honeypot | An attacker creates a seemingly vulnerable contract to lure users into believing that the money can be drained if a particular sum of funds are sent to the contract beforehand. The user's fund will be trapped, and can only be recovered by the attacker. | 1 | 2 | 474,236 |
| Darkweb Shop | Darkweb-related activity and/or fake illegal shops designed to steal funds. | 13 | 192 | 85,625 |
| Charity | Crypto projects impersonating charities after major events and asking for donations using phishing emails and websites. | 11 | 111 | 79,802 |
| Ransomware | A type of phishing scam whereby software is planted on the user's device in order to encrypt files. This can compromise crypto private key information as well as other credentials stored in the network unless a ransom is paid. | 7 | 36 | 77,836 |
| Hardfork Scam | Cybercriminals create a fake network upgrade of major blockchains and ask users to send respective tokens with the promise of new coins from the alleged new protocols. | 1 | 36 | 13,293 |
| Sextortion | Cybercriminals evoke fear by threatening victims with sharing their online behaviors such as visiting adult websites. Users are coerced into paying a ransom. | 4 | 70 | 7,475 |
| Other | | 137 | 4,181 | 4,974,454 |
| Total | | 5,644 | 1,783,981 | 1,651,407,924 |

**Note:** This table presents 19 cybercrime categories observed on the Ethereum blockchain. The list is derived from two primary sources: *Etherscan* and *Scam Alert*. *Etherscan*, a block explorer and analytics platform for Ethereum, assigns public name tags and labels to addresses that are of public interest. Any address associated with fraudulent activities has a brief warning message attached to it, providing investors with details of the purported scam. All blockchain addresses that *Etherscan* labeled as exploit, hack, heist, phish, Ponzi scheme, and/or scam are included. The authors have reclassified the scams into 19 finer categories based on the detailed information in the warning messages.

Table 2: **Matched variables and matching results**

| | Matching | Mean Treated | Mean Control | % bias | % bias reduction |
|---|---|---|---|---|---|
| Blockchain address age | Before | 7.6929 | 21.1131 | -164.41 | - |
| | After | 7.6929 | 7.6929 | 0.00 | 1.00 |
| Trading activity | Before | 4.2794 | 0.5966 | 7.18 | - |
| | After | 4.2794 | 4.1992 | 0.16 | 97.83 |
| Blockchain address balance | Before | 5,712.05 | 3,271.02 | 0.66 | - |
| | After | 5,712.05 | 5,266.08 | 0.12 | 81.73 |
| Diversification | Before | 1.6438 | 1.1025 | 18.25 | - |
| | After | 1.6438 | 1.6254 | 0.62 | 96.60 |

**Note:** This table presents mean values of matching variables for both the treatment (victims) and control (non-victims) groups. Each of our 200,865 victims is matched with a non-victim control with the lowest Euclidean distance score. These scores are determined based on blockchain address balance, blockchain address age, trading activity, and diversification, using data from the three months leading up to the public revelation of the scam. Our final sample comprises address–month observations from both our victim group (200,865) and our non-victim group (200,865). Definitions of all variables appear in Table A.1.




**Note:** Standardized % bias for each covariate is calculated as the difference in means in the treatment and control groups, divided by the standard deviation in the control group. This value is then represented as a percentage.

Table 3: **Summary statistics for victims of cybercrime (treatment group)**

| Variable | mean | stddev | min | max | q1 | median | q3 |
|---|---|---|---|---|---|---|---|
| return | 0.132 | 0.749 | -1 | 176.271 | -0.162 | 0 | 0.259 |
| churn rate | 0.055 | 0.358 | 0 | 59.656 | 0 | 0 | 0 |
| diversification | 2.346 | 6.408 | 1 | 637 | 1 | 1 | 1 |
| blockchain address balance | 9,218.949 | 3,145,094.447 | 0 | 2,826,598,945.187 | 2.52 | 15.702 | 75.566 |
| trading activity | 4.27 | 125.818 | 0 | 48,091 | 0 | 0 | 0 |
| blockchain address age (month) | 18.445 | 12.99 | 1 | 77 | 8 | 16 | 27 |
| lotterytoken investor (dummy) | 0.146 | 0.353 | 0 | 1 | 0 | 0 | 0 |
| lotterytoken share | 0.043 | 0.179 | 0 | 1 | 0 | 0 | 0 |
| stablecoin investor (dummy) | 0.048 | 0.214 | 0 | 1 | 0 | 0 | 0 |
| stablecoin share | 0.006 | 0.065 | 0 | 1 | 0 | 0 | 0 |
| altcoin share | 0.269 | 0.427 | 0 | 1 | 0 | 0 | 0.762 |
| ***3-factor model:*** | | | | | | | |
| diversifiable risk | 0.311 | 0.975 | 0 | 211.339 | 0.175 | 0.179 | 0.223 |
| non-diversifiable risk | 0.095 | 0.161 | 0 | 38.843 | 0.081 | 0.102 | 0.109 |
| total risk | 0.407 | 1.075 | 0 | 214.245 | 0.273 | 0.279 | 0.301 |
| market | 3.609 | 2.355 | -5.053 | 41.774 | 2.666 | 3.996 | 4.484 |
| momentum | -0.37 | 5.999 | -21.289 | 58.54 | -4.321 | -2.552 | 0.158 |
| size | 0.551 | 2.178 | -32.429 | 15.428 | 0.176 | 0.407 | 0.64 |
| alpha | 0.064 | 0.103 | -0.485 | 1.335 | 0.001 | 0.077 | 0.132 |

**Note:** This table reports summary statistics for victims of all fraud types. Variables are constructed monthly and our final sample includes address–month observations from 200,865 unique victim addresses. Definitions of all variables appear in Table A.1.

Table 4: **Summary statistics for victims (treatment group) and non-victims/non-cybercriminals (matched control group)**

| | | (1) | (2) | (3) | (4) |
|---|---|---|---|---|---|
| | | 3 months prior | | 3 months post | |
| Variable | Statistics | Victims | Non-victims | Victims | Non-victims |
| Return | mean | 0.090 | 0.085 | 0.214 | 0.147 |
| | (median) | (0.000) | (0.000) | (0.043) | (0.064) |
| Diversification | mean | 1.644 | 1.625 | 2.104 | 1.663 |
| | (median) | (1.000) | (1.000) | (1.000) | (1.000) |
| Blockchain address balance | mean | 5,712.046 | 5,266.077 | 8,178.883 | 6,686.706 |
| | (median) | (11.968) | (11.796) | (11.821) | (7.137) |
| Churn rate | mean | 0.088 | 0.09 | 0.081 | 0.071 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Trading activity | mean | 4.279 | 4.199 | 6.082 | 1.864 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Blockchain address age | mean | 7.693 | 7.693 | 10.345 | 10.345 |
| | (median) | (5.000) | (5.000) | (8.000) | (8.000) |
| Lottery token investor | mean | 0.107 | 0.127 | 0.118 | 0.127 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Lottery token share | mean | 0.028 | 0.041 | 0.034 | 0.038 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Stablecoin investor | mean | 0.070 | 0.129 | 0.075 | 0.136 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Stablecoin share | mean | 0.009 | 0.03 | 0.008 | 0.023 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| Altcoin share | mean | 0.098 | 0.111 | 0.225 | 0.137 |
| | (median) | (0.000) | (0.000) | (0.000) | (0.000) |
| ***3-factor model:*** | | | | | |
| Diversifiable Risk | mean | 0.105 | 0.106 | 0.301 | 0.218 |
| | (median) | (0.099) | (0.099) | (0.176) | (0.176) |
| Non-diversifiable Risk | mean | 0.199 | 0.195 | 0.117 | 0.108 |
| | (median) | (0.18) | (0.167) | (0.114) | (0.102) |
| Total Risk | mean | 0.305 | 0.300 | 0.418 | 0.325 |
| | (median) | (0.307) | (0.307) | (0.293) | (0.289) |
| Market | mean | 0.106 | 0.098 | 0.075 | 0.076 |
| | (median) | (0.000) | (0.000) | (0.088) | (0.088) |
| Momentum | mean | 2.584 | 2.464 | 3.374 | 3.608 |
| | (median) | (3.498) | (3.122) | (4.36) | (4.36) |
| Size | mean | 2.351 | 2.223 | 0.257 | -1.179 |
| | (median) | (0.000) | (0.000) | (-2.734) | (-2.283) |
| Alpha | mean | -1.93 | -1.773 | 0.569 | -0.104 |
| | (median) | (0.000) | (0.000) | (0.601) | (0.263) |

**Note:** This table presents the mean and median summary statistics for address–month observations for victims and matched non-victims for 3 months pre- (columns 1 & 2) and post-treatment (columns 3 & 4). Definitions of all variables appear in Table A.1.

Table 5: **Summary statistics for victims by scam type**

| Variable | Statistics | Ponzi Schemes | Giveaways | Phishing Scams | Investment Scams | Fake Token Sales | Hack/ Stolencrypto | Exploit/ Hardfork Scams | Darkweb Shop/ Exchange/ Charity | Sextortion Other |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | 3,174,851 | 149,386 | 529,074 | 55,961 | 21,477 | 1,807,369 | 18,849 | 139,386 | 63,363 |
| Return | Mean | 0.145 | 0.074 | 0.081 | 0.115 | 0.066 | 0.143 | 0.047 | 0.058 | 0.067 |
| | (Median) | (0.077) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Churn Rate | Mean | 0.032 | 0.141 | 0.168 | 0.286 | 0.25 | 0.027 | 0.401 | 0.266 | 0.145 |
| | (Median) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Diversification | Mean | 1.459 | 5.302 | 8.696 | 6.095 | 11.063 | 1.174 | 8.129 | 4.92 | 6.442 |
| | (Median) | (1.00) | (2.00) | (3.00) | (1.00) | (3.00) | (1.00) | (1.00) | (2.00) | (2.00) |
| Blockchain address Balance | Mean | 7,082.8 | 24,254.2 | 33,635.6 | 39,536.9 | 29,149.4 | 684.6 | 176,969.9 | 5,494.3 | 41,710.9 |
| | (Median) | (15.404) | (105.626) | (32.89) | (63.112) | (156.371) | (12.544) | (72.779) | (133.73) | (84.418) |
| Trading Activity | Mean | 1.235 | 28.519 | 13.967 | 67.978 | 12.847 | 1.002 | 27.323 | 11.556 | 16.671 |
| | (Median) | (0.00) | (0.00) | (0.00) | (1) | (0.00) | (0.00) | (1.00) | (0.00) | (0.00) |
| Blockchain address Age (Month) | Mean | 14.162 | 23.511 | 23.604 | 15.04 | 24.621 | 23.793 | 15.601 | 19.206 | 25.167 |
| | (Median) | (12.00) | (22.00) | (22.00) | (12.00) | (23.00) | (24.00) | (10.00) | (16.00) | (24.00) |
| Lotterytoken Investor (Dummy) | Mean | 0.063 | 0.575 | 0.674 | 0.43 | 0.731 | 0.042 | 0.575 | 0.414 | 0.608 |
| | (Median) | (0.00) | (1.00) | (1.00) | (0.00) | (1.00) | (0.00) | (1.00) | (0.00) | (1.00) |
| Lotterytoken Share | Mean | 0.016 | 0.156 | 0.23 | 0.116 | 0.203 | 0.014 | 0.133 | 0.065 | 0.148 |
| | (Median) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Stablecoin Investor (Dummy) | Mean | 0.038 | 0.092 | 0.109 | 0.3 | 0.115 | 0.006 | 0.421 | 0.352 | 0.073 |
| | (Median) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Stablecoin Share | Mean | 0.005 | 0.01 | 0.009 | 0.034 | 0.011 | 0.001 | 0.036 | 0.039 | 0.008 |
| | (Median) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Altcoin Share | Mean | 0.045 | 0.417 | 0.443 | 0.277 | 0.495 | 0.586 | 0.296 | 0.344 | 0.371 |
| | (Median) | (0.00) | (0.17) | (0.298) | (0.00) | (0.016) | (1.00) | (0.506) | (0.037) | (0.069) |

**Note:** This table presents the mean and median summary statistics for address–month observations for victims of various scam types. Definitions of all variables appear in Table A.1.

Table 6: **Post-scam changes in blockchain address-level risk-taking (treatment effects for victims vs. matched non-victims/non-cybercriminals), 3-month**

| | Average Treatment Effects | | Heterogeneous Treatment Effects by Cybercrime Type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | All scams (1) | All scams (2) | Ponzi scheme (3) | Give-away (4) | Phishing scam (5) | Investment (6) | Fake token scam (7) | Hack (8) | Exploit (9) | Darkweb shop (10) | Sextortion (11) |
| ***Panel A: Total Risk*** | | | | | | | | | | | |
| Post-scam | 0.0284*** | 0.0417 | -0.01*** | 0.015 | -0.004 | -0.015 | -0.01 | 0.122*** | 0.187** | -0.134 | -0.056** |
| | (0.002) | (0.003) | (0.002) | (0.018) | ( 0.012) | (0.033) | (0.014) | (0.004) | (0.066) | (0.082) | (0.018) |
| Victim | 0.1033*** | 0.1165*** | -0.014*** | 0.071*** | 0.13*** | 0.217*** | 0.035** | 0.454*** | 0.111⋆ | -0.161⋆ | -0.046** |
| | (0.002) | (0.003) | (0.002) | (0.014) | ( 0.018) | (0.067) | (0.013) | (0.005) | (0.043) | (0.076) | (0.015) |
| Post-scam × Victim | -0.0157*** | -0.0421*** | 0.021*** | -0.028 | -0.011 | -0.094 | 0.031 | -0.231*** | -0.345*** | 0.27 | 0.109*** |
| | (0.004) | (0.005) | (0.004) | (0.025) | ( 0.024) | (0.077) | (0.023) | (0.008) | (0.108) | (0.154) | (0.029) |
| Blockchain address age | 0.0033*** | 0.0035⋆ | 0.003*** | -0.001 | -0.001 | -0.002 | -0.001⋆ | 0.011*** | 0.001** | 0.004*** | 0.003*** |
| | (0.000) | (0.000) | (0.0) | (0.001) | ( 0.0) | (0.001) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.020 | 0.020 | 0.003 | 0.001 | 0.001 | 0.0 | 0.011 | 0.116 | 0.006 | 0.004 | 0.0 |
| No. obs. | 4,540,665 | 4,540,665 | 2,462,468 | 96,288 | 308,302 | 57,112 | 11,745 | 1,290,330 | 33,374 | 190,849 | 90,197 |
| ***Panel B: Non-Diversifiable Risk*** | | | | | | | | | | | |
| Post-scam | -0.0045*** | 0.0005*** | 0.001*** | -0.001 | -0.005*** | -0.025*** | -0.007 | 0.008*** | -0.011⋆ | -0.041*** | -0.034*** |
| | (0.000) | (0.000) | (0.000) | (0.002) | ( 0.001) | (0.006) | (0.004) | (0.001) | (0.006) | (0.009) | (0.006) |
| Victim | 0.0011*** | 0.0062*** | 0.011*** | 0.001 | -0.001 | 0.007 | -0.002 | -0.0 | -0.034*** | -0.044*** | -0.029*** |
| | (0.000) | (0.000) | (0.000) | (0.002) | ( 0.002) | (0.012) | (0.003) | (0.001) | (0.005) | (0.009) | (0.006) |
| Post-scam × Victim | 0.0043*** | -0.0058*** | -0.007*** | 0.005 | 0.01*** | 0.027 | 0.018** | -0.014*** | 0.031*** | 0.081*** | 0.067*** |
| | (0.001) | (0.001) | (0.001) | (0.003) | ( 0.002) | (0.014) | (0.006) | (0.002) | (0.009) | (0.018) | (0.012) |
| Blockchain address age | -0.0003*** | 0.0002*** | -0.004*** | -0.0 | -0.001*** | -0.0 | -0.001*** | -0.001*** | 0.001*** | -0.001*** | 0.0 |
| | (0.000) | (0.000) | (0.0) | (0.0) | ( 0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.007 | 0.014 | 0.011 | 0.028 | 0.01 | 0.002 | 0.05 | 0.03 | 0.052 | 0.028 | 0.037 |
| No. obs. | 4,540,665 | 4,540,665 | 2,462,468 | 96,288 | 308,302 | 57,112 | 11,745 | 1,290,330 | 33,374 | 190,849 | 90,197 |
| ***Panel C: Diversifiable Risk*** | | | | | | | | | | | |
| Post-scam | 0.0329*** | 0.0412*** | -0.011*** | 0.016 | 0.002 | 0.01 | -0.002 | 0.114*** | 0.198** | -0.094 | -0.022 |
| | (0.001) | (0.003) | (0.002) | (0.017) | ( 0.011) | (0.028) | (0.012) | (0.003) | (0.064) | (0.073) | (0.016) |
| Victim | 0.1022*** | 0.1103*** | -0.025*** | 0.07*** | 0.131*** | 0.21*** | 0.038*** | 0.454*** | 0.144*** | -0.117 | -0.017 |
| | (0.002) | (0.003) | (0.002) | (0.013) | ( 0.017) | (0.056) | (0.011) | (0.004) | (0.041) | (0.067) | (0.012) |
| Post-scam × Victim | –0.0200*** | –0.0363*** | 0.028*** | -0.032 | -0.021 | -0.121 | 0.013 | -0.217*** | -0.376*** | 0.189 | 0.042 |
| | (0.004) | (0.005) | (0.003) | (0.023) | ( 0.023) | (0.064) | (0.019) | (0.007) | (0.104) | (0.137) | (0.024) |
| Blockchain address age | 0.0037*** | 0.0036⋆ | 0.004*** | -0.0 | -0.001 | -0.001 | 0.0 | 0.01*** | 0.002*** | 0.003*** | 0.003*** |
| | (0.000) | (0.000) | (0.0) | (0.001) | ( 0.0) | (0.001) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.023 | 0.024 | 0.004 | 0.002 | 0.001 | 0.0 | 0.019 | 0.139 | 0.011 | 0.003 | 0.003 |
| No. obs. | 4,540,665 | 4,540,665 | 2,462,468 | 96,288 | 308,302 | 57,112 | 11,745 | 1,290,330 | 33,374 | 190,849 | 90,197 |

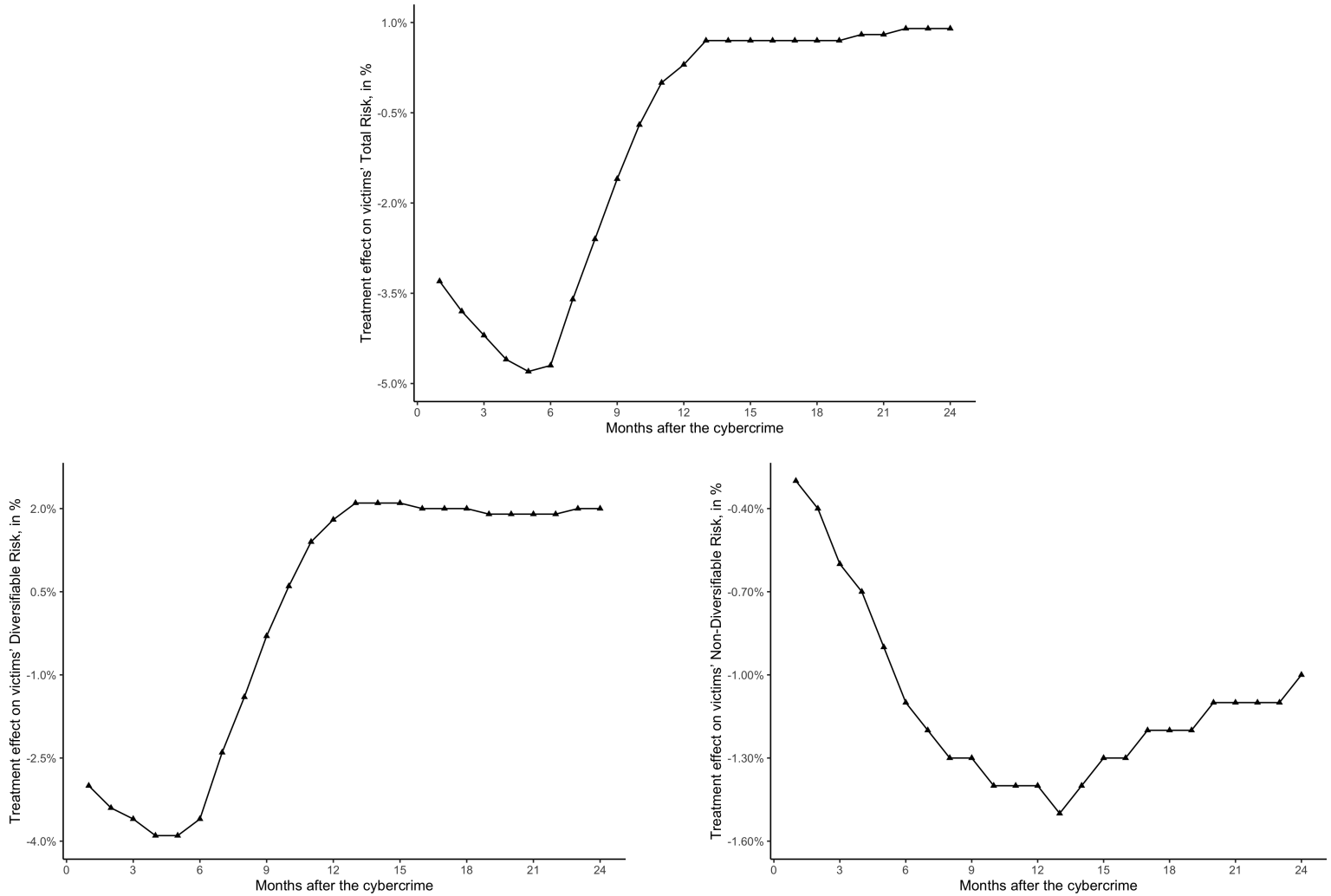
**Note:** These are difference-in-differences regressions to estimate the treatment effects of cybercrime on victim addresses' risk-taking. The dependent variables are total, diversifiable, and non-diversifiable risk-taking in Panels A, B, and C, respectively. The independent variables are a post-scam dummy that takes the value of 1 in the post-scam period, 0 otherwise; a victim dummy that takes a value of 1 for victims and 0 for matched non-victims, and their interaction term, i.e., the difference-in-differences estimator. We also control for blockchain address age, and include calendar-month and cybercrime-type fixed effects. The first two columns show the *average* treatment effects across all cybercrime types, while columns 3 to 11 show *heterogeneous* treatment effects for each cybercrime type separately. The regressions are estimated over the symmetric [–3, +3] event window with respect to the cybercrime month 0. Definitions of all variables appear in Table A.1.

Table 7: **Post-scam changes in blockchain address-level risk-taking (treatment effects for victims vs. matched non-victims/non-cybercriminals), 12-month**

| | Average Treatment Effects | | Heterogeneous Treatment Effects by Cybercrime Type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | All scams (1) | All scams (2) | Ponzi scheme (3) | Give-away (4) | Phishing scam (5) | Investment (6) | Fake token scam (7) | Hack (8) | Exploit (9) | Darkweb shop (10) | Sextortion (11) |
| ***Panel A: Total Risk*** | | | | | | | | | | | |
| Post-scam | 0.0058*** | 0.0158*** | –0.0314*** | 0.0287*** | 0.0130*** | 0.0475** | 0.0115 | 0.1109*** | 0.3913*** | –0.2327*** | –0.0617*** |
| | (0.001) | (0.001) | (0.001) | (0.012) | (0.008) | (0.023) | (0.009) | (0.003) | (0.092) | (0.052) | (0.012) |
| Victim | 0.0871*** | 0.0970*** | –0.0320*** | 0.0787*** | 0.1532*** | 0.2993*** | 0.0555*** | 0.4487*** | 0.2433*** | –0.2592*** | –0.0583*** |
| | (0.001) | (0.001) | (0.001) | (0.008) | (0.010) | (0.049) | (0.009) | (0.003) | (0.059) | (0.051) | (0.009) |
| Post-scam × Victim | 0.0230*** | 0.0033 | 0.0604*** | –0.0495*** | -0.0508 | –0.2390*** | –0.0109 | –0.2150*** | –0.6818*** | 0.4665*** | 0.1270*** |
| | (0.002) | (0.003) | (0.002) | (0.017) | (0.0014) | (0.064) | (0.014) | (0.006) | (0.150) | (0.100) | (0.019) |
| Blockchain address age | 0.0031*** | 0.0034*** | 0.0029*** | 0.000*** | –6.338e-05 | –0.0024*** | 3.465e-05 | 0.0134*** | 0.0003*** | 0.0033*** | 0.0030*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.001 | 0.018 | 0.002 | 0.000 | 0.001 | 0.000 | 0.009 | 0.117 | 0.009 | 0.004 | 0.001 |
| No. obs. | 7,837,125 | 7,837,125 | 4,661,547 | 152,595 | 520,739 | 84,237 | 19,133 | 1,985,775 | 35,379 | 268,411 | 109,309 |
| ***Panel B: Non-Diversifiable Risk*** | | | | | | | | | | | |
| Post-scam | –0.0019*** | 0.0039*** | 0.0069*** | –0.0032*** | –0.0084*** | –0.0274** | –8.606e-05 | 0.0082*** | 0.0036 | –0.0747*** | –0.0305*** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) | (0.004) | (0.002) | (0.000) | (0.005) | (0.007) | (0.003) |
| Victim | 0.0054*** | 0.0112*** | 0.0199*** | –0.0020** | –0.0030*** | 0.0082 | 0.0045** | –0.0010** | –0.0206*** | –0.0815*** | –0.0272*** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) | (0.009) | (0.002) | (0.000) | (0.004) | (0.007) | (0.003) |
| Post-scam × Victim | –0.0027*** | –0.0142*** | –0.0219*** | 0.0089*** | 0.0154*** | 0.0276** | 0.0035 | –0.0130*** | 0.0028*** | 0.1519*** | 0.0617*** |
| | (0.000) | (0.000) | (0.000) | (0.002) | (0.001) | (0.012) | (0.003) | (0.001) | (0.080) | (0.013) | (0.006) |
| Blockchain address age | –0.0001*** | 0.0000*** | –2.068e-05*** | –0.0005*** | –0.0003*** | –0.0012*** | –0.0008*** | 0.0006*** | –0.0013*** | 0.0002*** | 0.0003*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.003 | 0.008 | 0.012 | 0.015 | 0.005 | 0.001 | 0.029 | 0.020 | 0.049 | 0.013 | 0.033 |
| No. obs. | 7,837,125 | 7,837,125 | 4,661,547 | 152,595 | 520,739 | 84,237 | 19,133 | 1,985,775 | 35,379 | 268,411 | 109,309 |
| ***Panel C: Diversifiable Risk*** | | | | | | | | | | | |
| Post-scam | 0.0077*** | 0.0119*** | –0.0383*** | 0.0319** | 0.0214*** | 0.0115*** | 0.018 | 0.1027*** | 0.3877*** | –0.1580*** | –0.0312*** |
| | (0.001) | (0.001) | (0.001) | (0.011) | (0.007) | (0.020) | (0.008) | (0.002) | (0.089) | (0.046) | (0.011) |
| Victim | 0.0817*** | 0.0858*** | –0.0519*** | 0.0807*** | 0.1562*** | 0.2912 | 0.0510*** | 0.4497*** | 0.2639*** | –0.1776*** | –0.0311*** |
| | (0.001) | (0.001) | (0.001) | (0.008) | (0.010) | (0.041) | (0.007) | (0.003) | (0.057) | (0.046) | (0.008) |
| Post-scam × Victim | 0.0257*** | 0.0175*** | 0.0822*** | –0.0584*** | -0.0662*** | –0.2666** | –0.0144 | –0.2020*** | -0.6846*** | 0.3147*** | 0.0653*** |
| | (0.000) | (0.002) | (0.001) | (0.015) | (0.014) | (0.012) | (0.003) | (0.005) | (0.145) | (0.088) | (0.017) |
| Blockchain address age | 0.0032*** | 0.0034*** | 0.0029*** | 0.0006 | 0.0002 | –0.0013** | 0.0008*** | 0.0127*** | 0.0017*** | 0.0031*** | 0.0027*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.021 | 0.021 | 0.002 | 0.001 | 0.001 | 0.000 | 0.018 | 0.140 | 0.013 | 0.003 | 0.002 |
| No. obs. | 7,837,125 | 7,837,125 | 4,661,547 | 152,595 | 520,739 | 84,237 | 19,133 | 1,985,775 | 35,379 | 268,411 | 109,309 |

**Note:** These are difference-in-differences regressions to estimate the treatment effects of cybercrime on victim addresses' risk-taking. The dependent variables are total, diversifiable, and non-diversifiable risk-taking in Panels A, B, and C, respectively. The independent variables are a post-scam dummy that takes the value of 1 in the post-scam period, 0 otherwise; a victim dummy that takes a value of 1 for victims and 0 for matched non-victims, and their interaction term, i.e., the difference-in-differences estimator. We also control for blockchain address age, and include calendar-month and cybercrime-type fixed effects. The first two columns show the *average* treatment effects across all cybercrime types, while columns 3 to 11 show *heterogeneous* treatment effects for each cybercrime type separately. The regressions are estimated over the symmetric [–12, +12] event window with respect to the cybercrime month 0. Definitions of all variables appear in Table A.1.

Figure 3: **Changes in post-scam blockchain address-level risk-taking (treatment effects) over time**



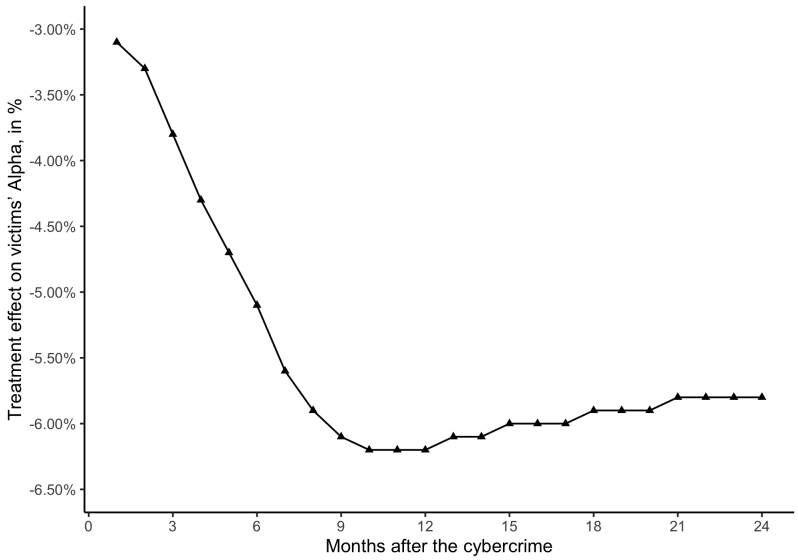

**Note:** These plots illustrate the time-structure of our identified treatment effects for victims' post-cybercrime risk-taking levels in terms of total risk (top), diversifiable risk (bottom left), and non-diversifiable risk (bottom right). The graphs plot the monthly coefficients from difference-in-differences models for the post-cybercrime months 1 to 24. Definitions of all variables appear in Table A.1.

Table 8: **Treatment effects of cybercrime on victims' risk-adjusted returns (alphas)**

| | Average Treatment Effects | | Heterogeneous Treatment Effects by Cybercrime Type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | All scams (1) | All scams (2) | Ponzi scheme (3) | Give-away (4) | Phishing scam (5) | Investment (6) | Fake token scam (7) | Hack (8) | Exploit (9) | Darkweb shop (10) | Sextortion (11) |
| ***Panel A: Alpha, 3-month event window*** | | | | | | | | | | | |
| Post-scam | 0.0156*** | 0.0171*** | 0.021*** | 0.002 | 0.007*** | -0.001 | -0.0 | 0.013*** | 0.101*** | -0.006 | -0.02*** |
| | (0.000) | (0.000) | (0.0) | (0.002) | ( 0.001) | (0.004) | (0.005) | (0.001) | (0.01) | (0.005) | (0.006) |
| Victim | 0.0148*** | 0.0163*** | 0.029*** | -0.001 | -0.001 | -0.013*** | -0.02*** | -0.011*** | 0.086*** | -0.018*** | -0.024*** |
| | (0.000) | (0.000) | (0.0) | (0.002) | ( 0.001) | (0.004) | (0.005) | (0.0) | (0.008) | (0.004) | (0.006) |
| Post-scam × Victim | –0.0353*** | –0.0383*** | -0.044*** | -0.001 | -0.01*** | 0.009 | 0.009 | -0.023*** | -0.201*** | 0.02⋆ | 0.039*** |
| | (0.000) | (0.001) | (0.001) | (0.003) | ( 0.001) | (0.006) | (0.009) | (0.001) | (0.017) | (0.009) | (0.011) |
| Blockchain address age | –0.0032*** | –0.0030*** | -0.004*** | -0.004*** | -0.002*** | -0.002*** | -0.003*** | -0.003*** | -0.001*** | -0.002*** | -0.003*** |
| | (0.000) | (0.000) | (0.0) | (0.0) | ( 0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.185 | 0.191 | 0.13 | 0.051 | 0.037 | 0.051 | 0.102 | 0.045 | 0.037 | 0.132 | 0.037 |
| No. obs. | 4,540,665 | 4,540,665 | 2,462,468 | 96,288 | 308,302 | 57,112 | 11,745 | 1,290,330 | 33,374 | 190,849 | 90,197 |
| ***Panel B: Alpha, 12-month event window*** | | | | | | | | | | | |
| Post-scam | 0.00156*** | 0.0285*** | 0.0380*** | 0.0014 | 0.0116*** | 0.0104*** | –0.0030 | 0.0130*** | 0.0996*** | –0.0091** | –0.0139*** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) | (0.003) | (0.003) | (0.000) | (0.011) | (0.004) | (0.003) |
| Victim | 0.0211*** | 0.0282*** | 0.0470*** | -0.0027** | 0.0033*** | –0.0044 | –0.0250*** | –0.0113*** | 0.0785*** | –0.0240*** | –0.0189*** |
| | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) | (0.003) | (0.003) | (0.000) | (0.008) | (0.004) | (0.003) |
| Post-scam × Victim | –0.0475*** | –0.0617*** | –0.0797*** | 0.0021 | –0.0186*** | –0.0120** | 0.0169*** | –0.0235*** | –0.1911*** | 0.0294*** | 0.0286*** |
| | (0.000) | (0.000) | (0.000) | (0.002) | (0.001) | (0.006) | (0.005) | (0.001) | (0.017) | (0.007) | (0.006) |
| Blockchain address age | –0.0031*** | –0.0031*** | –0.0039*** | –0.0022*** | –0.0015*** | –0.0024*** | –0.0026*** | –0.0012*** | –0.0018*** | –0.0026*** | –0.0021*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam-type FEs | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ |
| Adj. $R^2$ | 0.179 | 0.184 | 0.091 | 0.050 | 0.034 | 0.046 | 0.108 | 0.049 | 0.031 | 0.137 | 0.036 |
| No. obs. | 7,837,125 | 7,837,125 | 4,661,547 | 152,595 | 520,739 | 84,237 | 19,133 | 1,985,775 | 35,379 | 268,411 | 109,309 |

**Note:** These are difference-in-differences regressions to estimate the treatment effects of cybercrime on victims' risk-adjusted returns. The dependent variable are address-level alphas estimated from the three-factor crypto-asset pricing model in Liu et al. (2022). Panels A and B show regression results for the 3- and 12-month symmetric event windows, respectively. The independent variables are a post-scam dummy that takes the value of 1 in the post-scam period, 0 otherwise; a victim dummy that takes a value of 1 for victims and 0 for matched non-victims, and their interaction term, i.e., the difference-in-differences estimator. We also control for blockchain address age, and include calendar-month and cybercrime-type fixed effects. The first two columns show the *average* treatment effects across all cybercrime types, while columns 3 to 11 show *heterogeneous* treatment effects for each cybercrime type separately. Definitions of all variables appear in Table A.1.

Figure 4: **Post-cybercrime evolution of victims' blockchain address-level risk-adjusted returns**



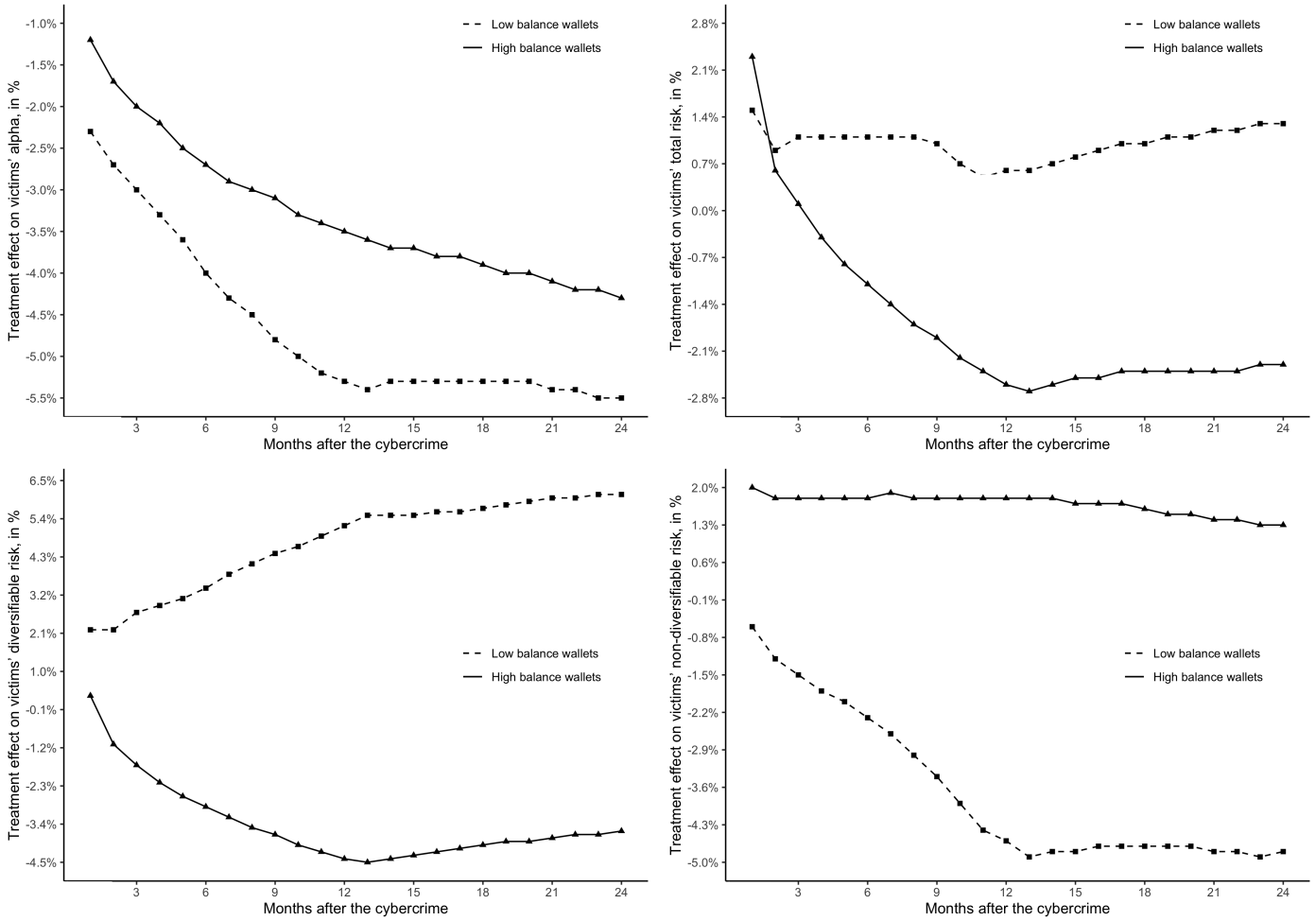

**Note:** This plot illustrates the time-structure of our identified treatment effect for victims' post-cybercrime risk-adjusted returns (i.e., alphas from the three-factor model introduced by Liu et al., 2022). The graphs plot the monthly coefficients from difference-in-differences models for the post-cybercrime months 1 to 24. Definitions of all variables appear in Table A.1.

Table 9: **Correlations between investor behavior, risk, and return in victims' post-scam blockchain addresses**

| | Trading activity (1) | Churn rate (2) | Diversification (3) | % Stablecoins (4) | % Altcoins (5) | % Lottery tokens (6) |
|---|---|---|---|---|---|---|
| ***Panel A: 3-month*** | | | | | | |
| Alpha × Post-scam × Victim | -20.067*** | -0.478*** | 1.517*** | 0.081*** | 0.164*** | 0.024 |
| | (3.543) | (0.083) | (0.309) | (0.005) | (0.048) | (0.014) |
| Alpha × Post-scam | 4.698*** | 0.151⋆ | -2.944*** | -0.033*** | -0.114*** | -0.066*** |
| | (1.207) | (0.068) | (0.163) | (0.003) | (0.02) | (0.009) |
| Alpha × Victim | -13.109*** | -0.032 | -5.912*** | -0.129*** | -0.421*** | 0.029*** |
| | (2.259) | (0.033) | (0.226) | (0.004) | (0.04) | (0.007) |
| Non-diversifiable risk × Post-scam × Victim | 2.142 | 0.083 | 0.626 | 0.083*** | 0.149 | -0.029 |
| | (3.714) | (0.088) | (0.615) | (0.013) | (0.113) | (0.022) |
| Non-diversifiable risk × Post-scam | -18.331*** | -0.028 | -1.079*** | -0.071*** | -0.137*** | 0.047** |
| | (2.28) | (0.075) | (0.318) | (0.009) | (0.039) | (0.016) |
| Non-diversifiable risk × Victim | 4.435⋆ | 0.009 | 1.025⋆ | -0.042*** | 0.125 | -0.003 |
| | (1.754) | (0.02) | (0.421) | (0.008) | (0.09) | (0.011) |
| Diversifiable risk × Post-scam × Victim | 0.097 | -0.022⋆ | -0.544*** | -0.009*** | -0.106*** | -0.024*** |
| | (0.316) | (0.009) | (0.118) | (0.002) | (0.022) | (0.006) |
| Diversifiable risk × Post-scam | 0.487⋆ | 0.024** | 0.508*** | 0.009*** | 0.06*** | 0.024*** |
| | (0.22) | (0.007) | (0.096) | (0.002) | (0.012) | (0.006) |
| Diversifiable risk × Victim | 0.216 | 0.016*** | 0.237*** | 0.005*** | 0.068*** | 0.01*** |
| | (0.185) | (0.005) | (0.061) | (0.001) | (0.018) | (0.003) |
| Post-scam × Victim | 6.612*** | 0.036*** | 0.345*** | -0.007*** | -0.007⋆ | -0.015*** |
| | (0.284) | (0.002) | (0.017) | (0.0) | (0.003) | (0.001) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam type FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Adj. $R^2$ | 0.007 | 0.065 | 0.128 | 0.031 | 0.229 | 0.066 |
| No. obs. | 4,540,665 | 4,540,665 | 4,540,665 | 4,540,665 | 4,540,665 | 4,540,665 |
| ***Panel B: 12-month*** | | | | | | |
| Alpha × Post-scam × Victim | -20.525*** | -0.132*** | 3.844*** | 0.111*** | 0.306*** | 0.018 |
| | ( 2.517) | ( 0.033) | ( 0.211) | ( 0.004) | ( 0.032) | ( 0.01) |
| Alpha × Post-scam | 2.998*** | -0.03 | -3.266*** | -0.033*** | -0.065*** | -0.008 |
| | ( 0.448) | ( 0.022) | ( 0.097) | ( 0.002) | ( 0.014) | ( 0.006) |
| Alpha × Victim | -11.687*** | -0.215*** | -7.623*** | -0.15*** | -0.577*** | -0.015*** |
| | ( 0.878) | ( 0.014) | ( 0.149) | ( 0.003) | ( 0.024) | ( 0.004) |
| Non-diversifiable risk × Post-scam × Victim | 3.397 | 0.175*** | 0.818 | 0.123*** | 0.507*** | 0.104*** |
| | ( 2.239) | ( 0.047) | ( 0.457) | ( 0.01) | ( 0.087) | ( 0.02) |
| Non-diversifiable risk × Post-scam | -14.604*** | -0.088** | -1.745*** | -0.101*** | -0.465*** | -0.055*** |
| | ( 0.933) | ( 0.032) | ( 0.243) | ( 0.006) | ( 0.041) | ( 0.014) |
| Non-diversifiable risk × Victim | 1.437** | -0.024 | 1.102*** | -0.052*** | 0.089 | -0.039*** |
| | ( 0.482) | ( 0.013) | ( 0.294) | ( 0.006) | ( 0.059) | ( 0.008) |
| Diversifiable risk : Post-scam × Victim | 0.239 | -0.022*** | -0.624*** | -0.014*** | -0.142*** | -0.04*** |
| | ( 0.153) | ( 0.005) | ( 0.077) | ( 0.001) | ( 0.015) | ( 0.005) |
| Diversifiable risk × Post-scam | 0.453*** | 0.025*** | 0.566*** | 0.011*** | 0.084*** | 0.032*** |
| | ( 0.087) | ( 0.003) | ( 0.06) | ( 0.001) | ( 0.01) | ( 0.004) |
| Diversifiable risk × Victim | 0.018 | 0.013*** | 0.269*** | 0.006*** | 0.078*** | 0.017*** |
| | ( 0.047) | ( 0.002) | ( 0.037) | ( 0.001) | ( 0.01) | ( 0.002) |
| Post-scam × Victim | 7.368*** | 0.057*** | 0.386*** | -0.004*** | 0.0 | -0.009*** |
| | ( 0.27) | ( 0.002) | ( 0.013) | ( 0.0) | ( 0.003) | ( 0.001) |
| Calendar-month FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Scam type FEs | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Adj. $R^2$ | 0.006 | 0.053 | 0.133 | 0.027 | 0.293 | 0.080 |
| No. obs. | 7,837,125 | 7,837,125 | 7,837,125 | 7,837,125 | 7,837,125 | 7,837,125 |

**Note:** These are difference-in-differences-in-differences (i.e., triple differences) regressions to explore correlations between various proxies for investor behavior and the estimated treatment effects for risk-adjusted returns and risk-taking. The dependent variables are trading activity, churn rate, diversification, stablecoin, altcoin, and lottery token holdings in % in columns 1, 2, 3, 4, 5, and 6, respectively. The independent variables are our difference-in-differences variables (victim dummy, post-scam dummy, and their interaction), which are simultaneously interacted with risk-adjusted returns, non-diversifiable risk-taking, and diversifiable risk-taking. We also include calendar-month and cybercrime-type fixed effects. The regressions are estimated over the symmetric [–3, +3] and [–12, +12] event windows with respect to the cybercrime month 0 in Panels A and B, respectively. Definitions of all variables appear in Table A.1.

Figure 5: **Heterogeneous treatment effects by blockchain address balance**




**Note:** These plots show heterogeneous treatment effects for affluent (top 10% richest addresses by pre-cybercrime balance) and non-affluent (bottom 10% poorest addresses) cybercrime victims. The outcome variables are victims' risk-adjusted returns (top left), total risk (top right), diversifiable risk (bottom left), and non-diversifiable risk (bottom right). The graphs plot the monthly coefficients from difference-in-differences models for the post-cybercrime months 1 to 24. Definitions of all variables appear in Table A.1.

Table 10: **Predictors of cybercriminals and victims**

| | | by Cybercrime Type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | All Scams (1) | Ponzi scam (2) | Give-away (3) | Phishing (4) | Investment scam (5) | Fake token (6) | Hack (7) | Exploit shop (8) | Darkweb (9) | Sextortion (10) |
| ***Panel A. Cybercriminals*** | | | | | | | | | | |
| Blockchain address age | -0.0 | 0.001* | -0.0 | 0.0 | 0.0 | 0.001 | -0.0 | -0.0 | 0.001 | 0.001 |
| | (0.0) | (0.001) | ( 0.0) | (0.0) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.0) |
| Diversification | 0.007*** | 0.01*** | 0.002*** | 0.008*** | 0.012*** | 0.017*** | 0.005*** | 0.009*** | 0.012*** | 0.017*** |
| | (0.0) | (0.001) | ( 0.001) | (0.0) | (0.002) | (0.005) | (0.001) | (0.001) | (0.001) | (0.001) |
| Lottery token share | -0.459*** | -0.534*** | -0.371*** | -0.49*** | -0.483*** | -0.538*** | -0.462*** | -0.367*** | -0.584*** | -0.623*** |
| | (0.003) | (0.018) | ( 0.007) | (0.005) | (0.014) | (0.033) | (0.014) | (0.03) | (0.021) | (0.011) |
| Stablecoin share | -0.331*** | -0.524*** | -0.403*** | -0.361*** | -0.344*** | -0.259*** | -0.247*** | -0.205*** | -0.302*** | -0.15*** |
| | (0.005) | (0.016) | ( 0.01) | (0.009) | (0.014) | (0.049) | (0.017) | (0.058) | (0.021) | (0.029) |
| Altcoin share | 0.202*** | 0.151*** | 0.19*** | 0.251*** | 0.016 | -0.135** | 0.155*** | -0.37*** | 0.121*** | 0.146*** |
| | (0.005) | (0.027) | ( 0.01) | (0.007) | (0.024) | (0.048) | (0.018) | (0.032) | (0.036) | (0.02) |
| Adj. $R^2$ | 0.138 | 0.212 | 0.081 | 0.177 | 0.152 | 0.156 | 0.211 | 0.141 | 0.187 | 0.164 |
| No. obs. | 82894 | 3580 | 24458 | 34714 | 6152 | 1182 | 4304 | 878 | 2498 | 5128 |
| ***Panel B. Victims*** | | | | | | | | | | |
| Blockchain address age | -0.004*** | -0.002*** | -0.001*** | -0.001*** | -0.001*** | -0.001*** | -0.006*** | 0.001*** | 0.0 | 0.001*** |
| | (0.0) | (0.0) | ( 0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Diversification | 0.01*** | 0.013*** | 0.011*** | 0.008*** | 0.006*** | 0.009*** | 0.008*** | 0.006*** | 0.01*** | 0.007*** |
| | (0.0) | (0.0) | ( 0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) |
| Lottery token share | -0.482*** | -0.486*** | -0.5*** | -0.447*** | -0.477*** | -0.48*** | -0.478*** | -0.561*** | -0.612*** | -0.587*** |
| | (0.0) | (0.001) | ( 0.002) | (0.001) | (0.003) | (0.004) | (0.001) | (0.004) | (0.001) | (0.002) |
| Stablecoin share | -0.472*** | -0.553*** | -0.378*** | -0.407*** | -0.438*** | -0.417*** | -0.265*** | -0.415*** | -0.352*** | -0.34*** |
| | (0.0) | (0.0) | ( 0.003) | (0.001) | (0.003) | (0.007) | (0.002) | (0.004) | (0.002) | (0.005) |
| Altcoin share | 0.293*** | 0.106*** | 0.226*** | 0.309*** | 0.262*** | 0.31*** | 0.533*** | 0.167*** | 0.191*** | 0.264*** |
| | (0.0) | (0.001) | ( 0.002) | (0.001) | (0.003) | (0.005) | (0.0) | (0.006) | (0.002) | (0.004) |
| Adj. $R^2$ | 0.134 | 0.106 | 0.219 | 0.216 | 0.265 | 0.256 | 0.245 | 0.369 | 0.332 | 0.269 |
| No. obs. | 13137045 | 7417854 | 314428 | 1143494 | 119771 | 48388 | 3643106 | 38171 | 284547 | 127286 |

**Note:** This table reports regression results examining the characteristics of addresses that belong to cybercriminals (Panel A) and victims (Panel B). The dependent variable is a dummy variable which takes a value of 1 if the address belongs to a cybercriminal (victim). Robust standard error in parenthesis. Definitions of all variables appear in Table A.1.

# Appendix

Table A.1: Variable definitions

| Variable | Definition |
|---|---|
| Barber and Odean (2000) raw return | Gross monthly return on investment using the beginning-of-day position statements. Following Barber and Odean (2000), all tokens are assumed to be bought or sold at the end of the month and ignore intra-month trading. The monthly return on the investor $i$'s portfolio is calculated as: $$return_{i,t} = \sum_{j \in Q} w_{j,t} R_{j,t}$$ where $j$ refers to different tokens in investor $i$'s portfolio at time $t$, $w$ refers to the weight of the $ value for the holdings of token $j$ in the total portfolio value at the beginning of the month, and $R$ is the gross monthly return of token $j$. |
| Total Risk | Total risk refers to the overall variability or volatility of the return for a specific address $i$ at time $t$. It encompasses all sources of risk, including both diversifiable and non-diversifiable components. |
| Diversifiable Risk | Diversifiable risk, also known as idiosyncratic risk, represents the portion of the total risk of the return that can be eliminated through diversification. It refers to the risk specific to the address $i$ at time $t$ and is captured by the term $\epsilon_{i,t}$ in the Fama-French 3-factor model. |
| Non-Diversifiable Risk | Non-diversifiable risk, also known as systematic risk, is the portion of the total risk of the return that cannot be eliminated through diversification. It captures the common risk factors that affect a broad range of addresses and is measured by subtracting the diversifiable risk from the total risk. |
| Churn rate | We measure investment horizon by calculating for each blockchain address how frequently the holder's positions are rotated on all of the portfolio's tokens (Gaspar et al., 2005). The churn rate of address $i$ at day $t$ is calculated as: $$churn\ rate_{i,t} = \frac{\sum_{j \in Q} \lvert N_{j,i,t} P_{j,i,t} - N_{j,i,t-1} P_{j,i,t-1} - N_{j,i,t-1} \Delta P_{j,t} \rvert}{\sum_{j \in Q} \frac{N_{j,i,t} P_{j,i,t} + N_{j,i,t-1} P_{j,i,t-1}}{2}}$$ where $P_{j,t}$ and $N_{j,i,t}$ represent the price and the number of tokens of token $j$ held by blockchain address $i$ at month $t$. |
| Blockchain address balance | We measure the balance of each address by summing over the $ value of all tokens held by a blockchain address: $$blockchain address\ balance_{i,t} = \sum_{j \in Q} N_{j,i,t} P_{j,i,t}$$ where $N_{j,i,t}$ and $P_{j,i,t}$ represent the number of tokens and the price of token $j$ held by address $i$ at month $t$. |
| Diversification | Diversification refers to the number of unique tokens held within an address at the end of each month. |

*(Continued)*

Table A.1 – Continued

| **Variable** | **Definition** |
|---|---|
| Trading activity | Trading activity represents the number of transactions, including purchases and sales, measured at the end of each month. |
| Blockchain address age | Blockchain address age denotes the duration in months since the address became active. |
| Lottery token investor | A dummy variable that takes a value of 1 if the investor has made investments in lottery tokens in that month. Lottery tokens are defined as tokens with a share price lower than 10 cents. |
| Lottery token share | Lottery token share corresponds to the proportion of the investor's total portfolio allocated to lottery tokens at the end of each month. |
| Stablecoin investor | A dummy variable that takes a value of 1 if the investor has made investments in stablecoins in that month. A token is deemed a stablecoin if it is designed to maintain a steady value, which can be achieved either by linking it to a specific commodity or currency, or by regulating its supply through algorithmic means. |
| Stablecoin share | Stablecoin share represents the percentage of the investor's total portfolio consisting of stablecoins at the end of each month. |
| Altcoin share | Altcoin share indicates the proportion of the investor's total portfolio allocated to altcoins at the end of each month. Altcoins are defined as tokens issued by start-ups to finance their blockchain projects. Currency tokens (ETH, WBTC, etc.) or stablecoins are excluded. |

Table A.2: **Cybercriminals (all types)**

| Variable | mean | stddev | min | max | q1 | median | q3 |
|---|---|---|---|---|---|---|---|
| return | 0.069 | 0.779 | -1 | 65.6 | -0.021 | 0 | 0.058 |
| churn rate | 0.076 | 0.38 | 0 | 18.151 | 0 | 0 | 0 |
| diversification | 4.521 | 12.431 | 1 | 287 | 1 | 1 | 2 |
| blockchain address balance | 56,246.866 | 2,072,803.766 | 0 | 229,506,948.254 | 0 | 0.29 | 39.413 |
| trading activity | 301.238 | 16,916.938 | 0 | 2,400,319 | 0 | 0 | 0 |
| blockchain address age (month) | 19.68 | 13.757 | 1 | 76 | 7 | 17 | 30 |
| lotterytoken investor (dummy) | 0.296 | 0.457 | 0 | 1 | 0 | 0 | 1 |
| lotterytoken share | 0.109 | 0.288 | 0 | 1 | 0 | 0 | 0 |
| stablecoin investor (dummy) | 0.106 | 0.308 | 0 | 1 | 0 | 0 | 0 |
| stablecoin share | 0.024 | 0.138 | 0 | 1 | 0 | 0 | 0 |
| altcoin share | 0.211 | 0.388 | 0 | 1 | 0 | 0 | 0.082 |
| ***3-factor model:*** | | | | | | | |
| diversifiable risk | 0.24 | 0.621 | 0 | 12.844 | 0.023 | 0.123 | 0.191 |
| non-diversifiable risk | 0.075 | 0.152 | 0 | 3.607 | 0.004 | 0.048 | 0.111 |
| total risk | 0.316 | 0.699 | 0 | 16.451 | 0.053 | 0.221 | 0.292 |
| market | 2.597 | 4.751 | -23.913 | 78.281 | 0.172 | 1.847 | 4.151 |
| momentum | 1.288 | 11.782 | -145.027 | 218.3 | -1.394 | 0 | 0.463 |
| size | -0.473 | 6.888 | -158.4 | 17.035 | -0.205 | 0.006 | 0.635 |
| alpha | 0.023 | 0.173 | -1.156 | 2.759 | -0.016 | 0 | 0.053 |

**Note:** This table reports summary statistics for cybercriminals of all fraud types. Variables are constructed monthly and our final sample includes address–month observations from 1,467 unique cybercriminal addresses. Definitions of all variables appear in Table A.1.

Table A.3: **Non-cybercriminals (all types)**

| Variable | mean | stddev | min | max | q1 | median | q3 |
|---|---|---|---|---|---|---|---|
| return | 0.061 | 0.843 | -1 | 89.998 | -0.075 | 0 | 0.077 |
| churn rate | 0.087 | 0.586 | 0 | 56.487 | 0 | 0 | 0 |
| diversification | 2.615 | 4.825 | 1 | 89 | 1 | 1 | 2 |
| blockchain address balance | 50,663.996 | 1,158,625.327 | 0 | 83,760,132.375 | 0 | 2.252 | 205.02 |
| trading activity | 23.416 | 1,199.593 | 0 | 147,241 | 0 | 0 | 0 |
| blockchain address age (month) | 19.682 | 13.756 | 1 | 76 | 7 | 17 | 30 |
| lotterytoken investor (dummy) | 0.318 | 0.466 | 0 | 1 | 0 | 0 | 1 |
| lotterytoken share | 0.092 | 0.264 | 0 | 1 | 0 | 0 | 0 |
| stablecoin investor (dummy) | 0.11 | 0.313 | 0 | 1 | 0 | 0 | 0 |
| stablecoin share | 0.016 | 0.115 | 0 | 1 | 0 | 0 | 0 |
| altcoin share | 0.208 | 0.381 | 0 | 1 | 0 | 0 | 0.079 |
| ***3-factor model:*** | | | | | | | |
| diversifiable risk | 0.221 | 0.653 | 0 | 13.644 | 0.018 | 0.158 | 0.202 |
| non-diversifiable risk | 0.085 | 0.137 | 0 | 2.287 | 0.009 | 0.062 | 0.114 |
| total risk | 0.306 | 0.713 | 0 | 14.066 | 0.079 | 0.234 | 0.299 |
| momentum | 0.834 | 8.988 | -79.788 | 123.634 | -1.663 | -0.002 | 0.436 |
| size | -0.368 | 6.606 | -138.591 | 29.716 | -0.438 | 0 | 0.722 |
| alpha | 0.025 | 0.265 | -0.696 | 7.623 | -0.026 | 0 | 0.054 |

**Note:** This table reports summary statistics for non-cybercriminals of all fraud types. Variables are constructed monthly and our final sample includes address–month observations from 1,467 unique non-cybercriminal addresses. Definitions of all variables appear in Table A.1.