BEYOND CODING:
OPEN-SOURCE DISCLOSURE AND STARTUP GROWTH

Wayne Landsman[*] Chen Lin[†] Daniel Rabetti[‡§] Che Zhang[¶]

August 2025

Abstract

We investigate how the disclosure of open-source code influences startup growth, focusing on the strategic trade-off between transparency-driven innovation and exposure to proprietary and cybersecurity risks. Using a novel dataset that links DeFi startups to GitHub activity, we find that open-sourcing significantly boosts user adoption and engagement, with effects that persist after controlling for market trends, protocol characteristics, and fixed effects. To address endogeneity, we implement a matched difference-in-differences design and use the staggered rollout of GitHub Copilot as an instrument. The results suggest a plausibly causal relation between open-source disclosure and startup growth, particularly for protocols with decentralized governance, venture backing, high-quality codebases, and active developer communities. However, open-source disclosure also carries risks: increased exposure to cyberattacks and competitive imitation. For example, after Uniswap's license expired, its market dominance fell by over 50% amid a surge in copycat platforms. Still, open-source projects exhibit stronger long-term innovation, developing more features after disclosure. We highlight a "transparency dilemma": while openness accelerates growth and innovation, it also heightens proprietary and security risks—a trade-off that startups must navigate as they go beyond coding to signal credibility and build success in open, highly collaborative markets.

JEL classification: G15, G18, G29, K29, K42, O16.
Keywords: Open-source innovation, code release, disclosure, decentralized finance, innovation management, startups.

[*] University of North Carolina, wayne_landsman@unc.edu.
[†] Hong Kong University (HKU) Busines School, chenlin1@hku.hk.
[‡] Corresponding Author: National University of Singapore (NUS) Business School, 15 Kent Ridge Drive, Singapore, 119245.
[§] Harvard Business School (visiting), drabetti@hbs.edu.
[¶] Tsinghua University, zhangche@sem.tsinghua.edu.cn.

The greatest cost of managerial short-termism is invisible: it is the innovation and growth that never happen because leadership was too focused on today's metrics to invest in tomorrow's possibilities.

– Adapted from Graham, Harvey, and Rajgopal (2005)

1 Introduction

The purpose of this study is to investigate how the disclosure of open-source code by decentralized finance (DeFi) startup platforms influences startup growth, with a particular focus on the strategic trade-off between transparency-driven innovation and exposure to proprietary and cybersecurity risks. Open-sourcing has become a powerful credibility signal in community-driven startup ecosystems (Lyandres, Palazzo, and Rabetti, 2022). By publicly releasing their source code, startups demonstrate a commitment to transparency and development, thereby fostering greater user trust and community engagement (Lerner and Tirole, 2005; Nagle, 2019; Gefen, Rabetti, Sun, and Zhang, 2024). In DeFi—an arena grounded in blockchain technology and open collaboration—transparency is especially valued as a foundation for accountability (Harvey and Rabetti, 2024). Such openness can help reduce information asymmetries between firms and capital providers, potentially easing financing frictions (Botosan and Plumlee, 2002; Francis, LaFond, Olsson, and Schipper, 2008; Howell, Niessner, and Yermack, 2020).

However, firms tend to disclose strategically when the informational benefits outweigh the risks of revealing sensitive information (Verrecchia, 1983; Dye, 1985; Healy and Palepu, 2001; Verrecchia, 2001). This trade-off is especially acute for technology startups, where intangible assets such as code are hard to value and easy to replicate (Aboody and Lev, 2000). Although disclosure may reduce information asymmetry, it also facilitates imitation and exploitation by rivals and adversaries (Bhattacharya and Ritter, 1983; Anton and Yao, 2002, 2004; Lerner and Schankerman, 2010, 2011). For example, Uniswap made its smart-contract code public to build community trust, but this allowed SushiSwap to copy the code and siphon away users in a "vampire attack."[1] Similarly, Compound's code was reused to launch Cream Finance (Cream Finance, 2020). The risks are not limited to imitation. In 2021, the Poly Network hack exploited a vulnerability in open smart contracts to steal over $600 million (Poly Network, 2021).[2] To

[1] Uniswap open-sourced its code to build trust and transparency, but this enabled SushiSwap to fork the code and lure users with token incentives—a strategy dubbed a "vampire attack" (Adams, 2020). The case highlights how openness can invite imitation and competitive entry.

[2] The Poly Network hack illustrates the risks of open-source code: public access to smart contract logic allowed attackers to spot and exploit a flaw in cross-chain verification, leading to one of the largest DeFi breaches (Poly Network, 2021).

date, over $6 billion have been hacked.[3]

To measure the trade-offs, we construct a comprehensive dataset linking startups' GitHub activity with on-chain outcomes. For each protocol, we match granular code release decisions to various outcomes including token prices, extent of adoption (measured by Total Value Locked, or TVL), governance structures, and investor types. The blockchain environment enables unusually precise measurement of outcomes: TVL reflects capital deposited by users in protocols, and token-level price reactions capture how investors interpret disclosure. Tokens are the native digital assets issued by protocols; they often provide governance rights or claims on future utility and can be traded on secondary markets (Amiram, Lyandres, and Rabetti, 2025). This distinction helps clarify that users contribute capital and liquidity to operate the platform, while investors (often tokenholders) bear the financial risks of disclosure events through immediate price reactions. Our dataset covers thousands of repository creation events, more than 100 million lines of code, and high-frequency financial and user engagement data. In this context, changes in TVL provide a standardized proxy for platform adoption and user engagement that is widely employed by both practitioners and academic researchers.[4]

We begin our empirical analysis by addressing our first research question: *does open-source disclosure spur startup growth?*

We start our empirical examination with an event-study framework that treats the launch of a new open-source repository as a quasi-exogenous information shock to the market and developer community. This approach provides high-frequency estimates of abnormal changes in TVL around disclosure dates, allowing us to trace short-run dynamics and rule out pre-trend bias. Consistent with disclosure acting as a meaningful signal, we find that protocols experience a +0.68 log-point surge in abnormal TVL within a week of disclosure, and the first open-sourcing yields +4.06 log points over the subsequent two weeks. These effects are also economically meaningful: given that the mean TVL per protocol in our sample is about $6 million, open-source disclosure generates, on average, an additional $4–5 million in economic activities within a week of disclosure.[5]

Transparency enabled community auditing—but also adversarial targeting.

[3] See https://defillama.com/hacks.

[4] See, e.g., Campello, Jin, Rabetti, and Saleh (2023); Coinmonks (2022); Cong, Prasad, and Rabetti (2025); Parlour (2023); Stablecoin Strategist (2023); Sygnum (2023); Messari (2023); De Simone, Jin, and Rabetti (2024); Cong, Rabetti, Wang, and Yan (2025); Landsman, Lyandres, Maydew, and Rabetti (2025); Saggese, Fröwis, Kitzler, Haslhofer, and Auer (2025).

[5] The mean TVL at disclosure is 15.62 in log scale, which corresponds to approximately $6.1 million in levels ($e^{15.62} \approx 6.1 \times 10^6$). A 0.68 log-point increase implies $e^{15.62+0.68} - e^{15.62} \approx 4.3$ million.

Next, we employ a fixed-effects panel regression design that leverages within-protocol variation in disclosure timing to examine the persistence of these adoption effects. Here, the dependent variable is daily abnormal log-TVL, and the key regressor is a post-disclosure indicator. This specification absorbs time-invariant heterogeneity through protocol or repository fixed effects, while controlling for common shocks via calendar-month, industry, and blockchain effects. The panel results confirm that disclosure-driven gains persist well beyond the event window, with usage expanding by roughly 10–15 percent relative to pre-event benchmarks and continuing to grow over time. Moreover, heterogeneity tests reveal stronger adoption responses among protocols with DAO governance, venture backing, and multi-chain deployment, consistent with external oversight and institutional support amplifying the credibility of transparency. Taken together, the event study and panel results provide evidence that open-source disclosure credibly signals development commitment and platform longevity, thereby reducing information asymmetries and spurring user engagement in settings where near-term risks are salient (Lerner and Tirole, 2005; Merton and Thakor, 2023; Harvey and Rabetti, 2024).

A critical challenge in interpreting our baseline results is the potential endogeneity of disclosure timing. Open-sourcing is a voluntary and often strategic decision, which raises two main endogeneity issues. The first is omitted variable bias: unobserved factors such as team quality, technological sophistication, or reputation may simultaneously increase both the likelihood of open-sourcing and the ability to attract users, thereby creating a spurious correlation between disclosure and growth. The second is reverse causality: protocols may choose to open-source in response to changes in adoption rather than as a driver of it. For example, a rapidly growing platform might release its code to signal quality and attract developers, while struggling teams may postpone disclosure to avoid scrutiny. In both cases, the observed link between open-source disclosure and startup growth could be mechanical rather than causal.

To strengthen the causal interpretation of our tests, we implement two identification strategies. First, we implement a matched difference-in-differences (DiD) design that compares protocols transitioning from closed to open source with observationally similar protocols that remain closed. Treated projects exhibit economically and statistically significant increases in abnormal TVL relative to matched controls, with adoption gains of about 16 percent within two weeks of disclosure. The effect emerges sharply after disclosure, persists in the subsequent weeks, and is not preceded by differential pre-trends, supporting a

causal interpretation.[6]

Second, we employ an instrumental variables (IV) design based on exogenous shocks to developer productivity: the staggered rollout of GitHub Copilot. Copilot—an AI-assisted coding tool—reduces the cost and time required to write, audit, and refine smart contract code. Its introduction (technical preview in June 2021, public release in mid-2022, and subsequent major updates) applied platform-wide across GitHub, was orthogonal to blockchain-specific conditions, but plausibly shifted resource-constrained DeFi teams toward releasing code earlier by alleviating development bottlenecks. The first-stage regression results reveal a 20–30 percent increase in the monthly probability of creating a new public repository in the months following Copilot updates, controlling for protocol and time effects. In the second stage, Copilot-induced open-source events lead to economically significant adoption gains, with TVL increases exceeding baseline OLS estimates.[7]

The positive impact of open-source disclosure on startup growth, contrasted with the fact that only 10.86% of startups in our sample with meaningful economic activity choose to disclose their source code,[8] naturally raises a follow-up research question: *why do so few startups choose to open-source?*

Despite its benefits, openness also exposes protocols to imitation and security risks that can erode first-mover advantages. The expiration of Uniswap's V3 code license in April 2023 exemplifies this dynamic: dozens of forks emerged, industry concentration declined sharply, and liquidity became fragmented across platforms. Our difference-in-differences analysis shows that decentralized exchange (DEX) protocols directly affected by this license expiration experienced significantly altered adoption dynamics relative to other DeFi segments, which is consistent with a surge in copycat competition. Cybersecurity risks further amplify these costs: in our sample, open-source adopters face a 5–6 percentage point higher probability of a major exploit in the year following disclosure, with average dollar losses roughly 8 percent larger than those of their closed-source peers.

These proprietary and security concerns are also reflected in the reactions of another key audience exposed to disclosure: token holders. Token prices decline following disclosure—by approximately –1.6 percent over a two-day window—with no evidence of anticipatory run-ups. The immediate sell-off

[6] We match protocols on observable characteristics such as size, age, governance structure, and blockchain affiliation, and implement dynamic event-time plots to confirm parallel pre-trends. Inferences are robust to alternative matching procedures and sample restrictions.

[7] Landsman et al. (2025) employs a similar design, introducing ChatGPT as an IV that affects the probability of DeFi protocols choosing a centralized auditor.

[8] We filter out startups with less than a million in TVL to distinguish between failed startups and those that succeeded but chose to keep their source code closed.

suggests that short-horizon traders interpret disclosure as exposing firms to competitive imitation and heightened attack surfaces (Lerner and Schankerman, 2011; Ellis, Fee, and Thomas, 2012).[9] These short-run market penalties underscore the central trade-off: founders must weigh the adoption benefits of transparency against the possibility of undermining stakeholders' confidence, particularly when token prices are sensitive to short-term risks.[10]

Over longer horizons, token values recover and eventually exceed pre-disclosure levels—+7.9 percent at 90 days and +17.9 percent at six months—as markets begin to recognize the benefits of open-source development. This gradual revaluation reflects a compositional shift in the investor base: short-term speculators exit, while long-horizon investors realign around the platform's strengthening fundamentals. Such dynamics are consistent with theories of investor underreaction to innovation (Cohen, Diether, and Malloy, 2013) and with noise introduced by "code-washing" practices, where startups may release superficial code to feign innovation (Gefen et al., 2024). Ultimately, markets reward the deeper network effects, developer participation, and collaborative innovation that open-sourcing enables, aligning valuations with long-term value creation (Cong, Li, and Wang, 2022).

Having established evidence on the trade-offs of open-source disclosure, we provide three mechanisms that help explain why managers choose to release source code despite heightened risks of imitation and cybersecurity breaches in unregulated DeFi markets. First, there is a governance and monitoring mechanism: transparency is most credible and impactful when accompanied by strong oversight, as indicated by significantly larger adoption gains for protocols governed by decentralized autonomous organizations (DAOs) or backed by venture capital. This pattern is consistent with external monitors amplifying the credibility of disclosure in settings where formal regulation is weak (Lyandres et al., 2022; Cong et al., 2025). Second, there is an innovation-signaling mechanism: the informational content of a disclosure matters. Repositories with greater technological salience—a reflected by proxies including GitHub stars, forks, or transparency-focused modules—generate stronger user responses, much like how R&D announcements or patents signal quality in traditional industries (Eberhart, Maxwell, and Sid-

[9] In contrast, Emery, Lim, and Ye (2024) documents that repository releases by U.S. listed firms generate only a modest cumulative abnormal return (CAR) of 0.12 percent over a three-day window, suggesting that institutional investors may be better positioned to diversify risk and value long-run benefits.

[10] If founders anticipate short-term price declines, those catering to near-term token sales or short-horizon investors may delay or avoid open-sourcing, even when the long-run NPV of disclosure is positive. This dynamic mirrors survey evidence from Graham, Harvey, and Rajgopal (2005), where executives admitted to forgoing valuable long-term projects to avoid missing quarterly targets, and is consistent with models of short-termist behavior in capital markets (Stein, 1989; Bhojraj, Hribar, Picconi, and McInnis, 2009).

dique, 2004).[11] Third, there is a developer-community engagement mechanism: protocols with larger and more interconnected developer networks realize stronger adoption gains from open-sourcing, reflecting the role of collaborative innovation in the DeFi ecosystems (Cong, Li, and Wang, 2021).

This paper makes three primary contributions. First, we extend classic voluntary disclosure theory (Verrecchia, 1983; Dye, 1985; Verrecchia, 2001) to open-sourcing disclosure in highly innovative startups. Disclosing source code can greatly amplify user adoption and collaborative innovation, but it comes at the cost of proprietary information loss and heightened security vulnerabilities. In line with prior work noting that open-source and proprietary strategies entail different costs and benefits (Lerner and Schankerman, 2011), our findings highlight that entrepreneurial firms must balance transparency against potential attacks or IP expropriation when deciding to open-source their code. This mirrors broader concerns in the accounting literature on how firms strategically weigh disclosure benefits against proprietary costs, especially when revealing information could enable imitation or competitive harm (e.g., Ellis et al. (2012)). Overall, our study extend voluntary disclosure theory to a new frontier, showing how managerial incentives and investor biases interact with cybersecurity and imitation costs in shaping disclosure choices and economic outcomes in a newly unorthodox and highly innovative market (Leuz and Wysocki, 2016).

Second, and most germane to the innovation literature in accounting, we address recent calls to move beyond the predominant focus on patent disclosures and publicly listed firms (Barth and Gee, 2024; Glaeser and Lang, 2024). The vast majority of existing studies examine patent filings, R&D expenditures, or intellectual-property disclosures in corporate settings (e.g., Glaeser and Landsman (2021); Allen, Lewis-Western, and Valentine (2022); Beyhaghi, Khashabi, and Mohammadi (2023); Hegde, Herkenhoff, and Zhu (2023); Dyer, Glaeser, Lang, and Sprecher (2024)). Yet regulators and scholars increasingly emphasize that innovation is a broader concept whose information challenges extend well beyond patents and large public firms.[12] In our setting, innovation outputs (software code) are made public voluntarily, allowing us to observe how transparency in innovation unfolds when a firm's core technology is openly shared in real time rather than disclosed through traditional patent filings or financial statements. We show that such voluntary code releases yield significant knowledge spillovers and collaborative benefits—

[11] A GitHub "star" is an endorsement that allows users to bookmark a project; a "fork" is a copy of a repository that developers can modify for their own use; and transparency-focused modules are code components designed to improve auditability and monitoring, such as tools for logging, verification, or governance tracking.

[12] Fewer than 44% of Swedish innovations between 1970–2015 were ever patented and patent data capture only about 17% of innovation information (Taalbi (2022)). The Federal Trade Commission emphasizes that patents can both under- and overrepresent innovation activity, and calls for complementary measures (https://tinyurl.com/rahv8k2p).

echoing the positive externalities of innovation disclosure noted in prior studies—yet also impose proprietary costs on the disclosing firm, much like the trade-offs documented in corporate settings. By providing evidence on these dynamics in an extreme and underexplored setting, our study helps bridge the gap between traditional accounting research on innovation and the realities of modern, open-collaboration business models, thereby shedding light on how accounting information systems and disclosure practices can evolve to better capture value creation in emerging digital ecosystems.

Third, we contribute to the literature on the misvaluation of innovation. Prior research shows that markets often fail to correctly price innovative activities, such as the persistent underreaction to R&D disclosures (Cohen et al., 2013). This is particularly notable given that openness can generate substantial long-term gains; for example, recent evidence suggests that startup engagement with open-source communities enhances funding success (e.g., Bourveau, De George, Ellahie, and Macciocchi (2022); Lyandres et al. (2022); Conti, Peukert, and Roche (2024)). However, in settings characterized by high information asymmetry and slow verification, transparency can also distort signals. For instance, "code-washing" tactics demonstrate how ventures may strategically manipulate disclosure to mislead stakeholders during fundraising (Gefen et al., 2024). We extend this literature by showing that capital markets may initially misprice the long-term benefits of open-source projects, as investors discount their positive NPV due to concerns over imitation and cybersecurity risks.

Our findings also carry implications for startup founders and the governance of DeFi platforms (Cong et al., 2025). Entrepreneurs in high-tech sectors must strategically navigate transparency decisions—balancing the community's demand for openness against the imperative to protect their innovations from exploitation. Meanwhile, investors and community governors should recognize the long-term value generated by transparency and possibly design mechanisms—such as bug bounties (Landsman et al., 2025), insurance, or staged disclosure strategies—to mitigate potential short-run risks.

The remainder of the paper proceeds as follows. Section 2 describes the data. Section 3 presents the baseline results. Section 4 discusses identification strategies aimed at mitigating endogeneity and presents DiD and IV results. Section 5 examines disclosure costs and long-run implications. Section 6 explores potential economic mechanisms. Section 7 concludes.

2 Methodology and Data

2.1 Literature Review

Classic voluntary disclosure theories posit that firms reveal information selectively, balancing the benefits of transparency against proprietary costs (Verrecchia, 1983). In seminal work, (Verrecchia, 1983) formalizes managers' incentives to withhold sensitive information unless the expected benefits of disclosure outweigh the competitive risks, a trade-off further developed by (Dye, 1985) and others (Healy and Palepu, 2001; Verrecchia, 2001). This tension is especially pronounced for technology startups holding intangible assets—such as software code—that are difficult to value yet easy for rivals to appropriate (Aboody and Lev, 2000). Consistent with these theories, empirical evidence shows that firms often avoid divulging commercially sensitive details when proprietary costs are high. For instance, companies are more reluctant to disclose information about major customers if such disclosure could aid competitors, reflecting the real economic stakes of transparency (Ellis et al., 2012). In short, greater disclosure can reduce information asymmetry, but it also risks eroding a firm's competitive advantage and intellectual property.

A growing stream of literature examines these trade-offs in the context of open-source innovation. (Lerner and Tirole, 2005) discuss the economics of technology sharing, highlighting how open-source participation can serve as a credibility signal and development strategy for entrepreneurial firms. Subsequent studies find that contributing to open-source software projects is associated with higher firm productivity (Nagle, 2019) and can enhance startups' innovation capabilities as well as access to financing (Lyandres et al., 2022). Engaging with developer communities by publicly releasing code appears to build reputational capital: startups active on GitHub tend to attract more users and investors. At the same time, openness entails significant risks. Prior work notes that open-source versus proprietary models involve different cost-benefit profiles (Lerner and Schankerman, 2011), and cases in practice illustrate how releasing code can invite imitation and cyberattacks. For example, Uniswap's decision to open-source its smart-contract code helped build trust but also enabled a competitor to copy the code and siphon away users in a "vampire attack," undermining Uniswap's market position. Consistent with such concerns, some ventures may only superficially open their code to appear transparent without truly revealing valuable IP—a practice dubbed "code-washing" (Gefen et al., 2024). Thus, the decision to go open-source represents a strategic disclosure choice: it can foster collaboration and user growth, but it also heightens

exposure to imitation and security breaches.

Transparency can yield important benefits by alleviating information asymmetries, especially in capital markets. A robust body of research links higher disclosure quality to lower financing frictions: firms with more extensive voluntary disclosure enjoy a reduced cost of equity capital (Botosan and Plumlee, 2002), and better earnings transparency is associated with lower information risk for investors (Francis et al., 2008). In emerging fundraising settings such as cryptocurrency token offerings, credible voluntary disclosure similarly underpins investor trust and project success (Howell et al., 2020). However, even when innovation information is disclosed, markets do not always price it efficiently. Studies in accounting and finance find that investors often undervalue innovative investments and other intangibles in the short run, leading to predictable excess returns when the long-term benefits of those innovations materialize. For example, firms that substantially increase their R&D expenditures tend to earn positive abnormal stock returns subsequently, consistent with initial under-reaction to innovation news (Eberhart et al., 2004). Similarly, markets exhibit persistent underreaction to announcements of R&D breakthroughs, mispricing the future gains from innovation (Cohen et al., 2013). This evidence of misvaluing innovation suggests that transparency alone may not immediately eliminate information frictions in innovative settings - a consideration that is highly relevant to understanding how open-source disclosures are received by users and investors.

An emerging stream of accounting literature focuses explicitly on innovation and its unique information challenges. This literature has expanded dramatically in recent years: of nearly 200 accounting studies on innovation surveyed by Glaeser and Lang (2024), more than half were published in the last five years. A defining theme in this line of research is that innovation differs from other business activities due to its novelty, nonrivalry, and partial excludability. These attributes help innovation drive economic growth, but they also create unique information problems that accounting researchers are well-suited to address. For example, because innovations are nonrival (many parties can use the idea simultaneously) yet only partially excludable, the act of disclosure generates positive knowledge spillovers to other firms and society (Barth and Gee, 2024). At the same time, disclosure imposes proprietary costs on the innovator by enabling imitation, potentially dampening the original firm's incentive to invest in R&D (Verrecchia, 1983; Kim and Valentine, 2021). This tension between the external benefits and private costs of innovation disclosure is analogous to the classic disclosure trade-off in accounting (e.g., Verrecchia, 1983) and is borne out in empirical findings: for instance, when patent laws forced firms to disclose inventions earlier

and more fully, firms subsequently curtailed their own R&D activity, consistent with proprietary cost concerns (Kim and Valentine, 2021).

Given these challenges, accounting scholars have examined a variety of mechanisms to encourage innovation while protecting innovators. Studies have investigated how financial reporting rules and regulation impact innovation investment (e.g., the effects of Sarbanes–Oxley on R&D spending; Gao, Wu, and Zimmerman, 2017), how managers strategically time and format innovation disclosures (e.g., publishing patent applications or redacting trade secrets in filings; Heeley, Matusik, and Jain, 2007; Heeley and Jacobson, 2008; Glaeser, Michels, and Verrecchia, 2020; Glaeser and Landsman, 2021; Valentine, Zhang, and Zheng, 2025), and how firms use tax planning and contracting to support innovation(Bloom, Schankerman, and Van Reenen, 2013; Plumlee, Xie, Yan, and Yu, 2015; Cheng, Guo, Weng, and Wu, 2021. A key insight is that accounting information can mitigate certain innovation frictions—such as managerial myopia or capital misallocation—yet standard accounting measures struggle to capture the value of innovative intangibles.

Accordingly, a critical focus in recent work is the measurement of innovation. Researchers have employed diverse proxies (R&D expenditures, patent counts and citations, textual disclosures, etc.), each with strengths and pitfalls. Glaeser and Lang (2024) caution that no single metric fully captures innovation output, and they highlight common pitfalls (e.g., truncation biases in patent data, misclassification of R&D) that can lead to measurement error. They advocate for developing new empirical techniques, novel datasets, and creative sources of exogenous variation to better measure innovation and to identify causal relationships. In summary, the accounting literature on innovation is converging around the idea that innovation's unique economic characteristics necessitate tailored reporting and disclosure approaches. Ongoing research opportunities lie in embedding innovation more deeply into accounting frameworks and examining how accounting can support innovation while preserving creator incentives. Our study addresses key gaps in the literature by analyzing an unconventional disclosure form—open-source code release—and its economic consequences for consumers, entrepreneurs, and investors.

2.2 Data and Sample Selection

Our primary sample consists of startups in the decentralized finance (DeFi) sector that launched between 2019 and 2024. We define a startup as a venture developing a blockchain-based financial application

(such as a lending protocol, decentralized exchange, asset management platform, etc.) that typically issues its own crypto token and smart contract platform (Lyandres et al., 2022).[13] To construct the sample, we began with a comprehensive list of DeFi projects aggregated from industry databases (including DeFi Pulse and DeFi Llama), blockchain developer forums, and venture capital reports. We then applied three selection criteria: (1) the project must have an identifiable GitHub organization or repository for its code, (2) the project's launch and development activity must fall within our 2019–2024 observation window, and (3) the GitHub repository launch should occur after the protocol's initial launch, to ensure that post-disclosure dynamics (e.g., TVL) are not confounded by pre-existing activity. This process yielded 263 unique DeFi protocols that have at least one repository on the GitHub platform, and 3,900 unique repositories that disclose coding activities to the community.

Figure 1 illustrates the monthly number of new open-source repositories launched by DeFi protocols between 2019 and 2024. Repository creation surged sharply beginning in early 2021, reaching multiple peaks between mid-2022 and early 2023. This period of heightened activity was followed by a gradual yet volatile decline throughout the remainder of 2023 and into 2024.

[Figure 1 here]

For each defi protocol, we collected detailed data on coding activity and economic outcomes. The coding data comes from GitHub's API and public archives, from which we extract the universe of software repositories associated with each venture, along with time-stamped events such as repository creations (initiations), code commits, forks, and stars. In total, our dataset includes over 10,000 repositories and millions of individual code commits or other GitHub events. We use these data to construct measures of open-source disclosure. Our primary indicator is a dummy variable, OpenSourceLaunch, which equals 1 on the date a DeFi protocol publicly releases a new GitHub repository (and remains 1 thereafter in panel regressions). In additional analyses, we focus specifically on the first open-source repository launch–i.e., when a previously closed-source protocol publishes its first public repository–thereby identifying the effect of transitioning from closed to open source.

We obtain daily economic performance measures from multiple sources. The Total Value Locked (TVL) for each platform (in USD) is sourced from DeFi analytics platforms, which track the value of

[13] See Luo, Rabetti, and Yu (2024) and Lee, Pinto, Rabetti, and Sadka (2024) for blockchain applications in the corporate setting.

assets deposited in DeFi protocols' smart contracts. TVL is a key gauge of user engagement and platform traction in this industry. We also retrieve daily token prices and market capitalizations from cryptocurrency market databases (such as CoinGecko and CoinMarketCap). For startups without a tradable token, we use other proxies for value, such as the amount of capital raised or valuation implied in funding rounds, although our main analyses focus on the subsample with tokens to capture market-based reactions. Additionally, we collect various control variables: indicators for major news events or security incidents (e.g., hacks), Bitcoin market returns (to control for general crypto market movements), and on-chain usage metrics like number of active users or transaction volumes on the platform.

We include a series of DeFi protocol-level and repository-level characteristics. The main variables are described as below. *Size* is defined as the natural logarithm of a protocol's Total Value Locked (TVL) as of the GitHub repository's creation date. *AUDIT* is an indicator variable equal to one if a DeFi protocol's smart contracts have been audited by at least one auditor. *ORACLE* is an indicator variable equal to one if a protocol integrates external data feeds (e.g., asset prices) via oracle infrastructure. *LISTED* is an indicator variable that equals to one if the protocol has issued at least one tradable token on an exchange. *DAO* is an indicator variable equal to one if the protocol is governed by a Decentralized Autonomous Organization (DAO). *Lending* is an indicator variable that equals to one if a protocol whose primary functionality centers around lending and borrowing services. *Venture capital* is an indicator variable equal to one if the protocol used to receive early-stage venture capitcal(VC) funding. *HACK* is an indicator variable that equals to one if the repository launch month coincides with one or more documented DeFi security breaches. *Twitter followers (log)* is defined as the log-transformed number of followers of the protocol's official Twitter account. *Chains (log)* is defined as the natural logarithm of blockchains on which a protocol's smart contracts are deployed. *Staking (log)* is defined as the natural logarithm of assets staked within the protocol, capturing user engagement in governance or network security.

In repository-level, *Popularity* is defined as the repository's stargazer count on GitHub. *Extensibility* is measured by the number of forks a repository received, capturing downstream usage or adaptation of the coding. *Transparency* is a binary indicator equal to 1 if the repository is identified as analytics-oriented, based on textual analysis of metadata fields, identifying references to data reporting, dashboard functionality, or user-facing metrics. *Developer commit productivity* is defined as the average number of code commits by a repository's developers. *Developer team size* is defined as the number of distinct developers associated with the protocol before the repository launch. *Developer network* is defined as the

average number of DeFi platforms each developer is affiliated with at the time of launch, reflecting cross-platform expertise.

Table 1 presents summary statistics for the sample. Panel A documents the characteristics of DeFi ventures at the time they launch their open-source repository. On average, these protocols exhibit a total value locked (TVL) of 15.62 in logarithmic scale. Approximately 74% of the ventures have undergone at least one formal smart contract audit. Around 27% of protocols integrate off-chain data via oracles, with an average log staking value of 4.23, reflecting significant user participation through locked assets. The lending platforms account for 12% of the sample. While 16% of the protocols operate under decentralized autonomous organizations (DAOs), 2.2% are exposed to high-risky hack events. On the fundraising performance, 25% of ventures are backed by venture capital, whereas fewer than half (47%) have listed tokens on public exchanges. On the social and technical dimensions, the average venture has over 7,000 Twitter followers (log), is deployed across approximately 1.13 blockchains (log).

Panel B reports statistics on open-source repositories affiliated with these ventures. Repository popularity varies widely, with an average of 4.69 stargazers, suggesting that only a small subset gain substantial visibility. Extensibility metrics, proxied by the number of forks, similarly average 4.80. Approximately 4% of repositories are classified as analytics-oriented, based on their metadata characteristics.[14] The typical developer team consists of fewer than two contributors. In terms of productivity, individual developer generate an average of 792 commits before. Importantly, developers demonstrate cross-platform engagement: each one is engaged, on average, with over five distinct DeFi protocols, highlighting the potential for knowledge diffusion.

[Table 1 here]

2.3 Empirical Strategy

Our empirical analysis proceeds in two parts: first, an event study examining user responses to open-source disclosures; second, a panel regression framework estimating the impact of open-sourcing on startup growth (i.e., total value locked (TVL)), while controlling for a rich set of protocol and repository characteristics.

[14] We identify repositories as *analytics-oriented* based on keyword matching in their metadata descriptions, including terms such as "dashboard," "reporting," "metrics," "analytics," and "visualization." This classification aims to capture codebases primarily intended to facilitate user understanding of the protocol's underlying technology and operational metrics.

2.3.1 Event Study

We employ an event study framework to assess the effect of open-source disclosures on user engagement, measured by Total Value Locked *(TVL)*. We treat the public repository launch by a DeFi protocol as a quasi-exogenous information shock to the market and broader developer community. The event date, denoted as $t = 0$, corresponds to the OpenSourceLaunch date, and we examine the evolution of outcome variables in event time around this benchmark.

To capture market responses to open-source code disclosure, we implement a cumulative abnormal Total Value Locked (*CATVL*) framework following the methodology of Amiram, Jørgensen, and Rabetti (2022). We compute the Abnormal TVL (*AbTVL*), defined as the deviation of a protocol's actual daily TVL from a pre-event benchmark. TVL is widely used to capture economic activity in the DeFi literature (e.g., our primary outcome variable, Total Value Locked (*TVL*), represents the total dollar value of all assets deposited in a protocol (e.g., (Cong et al., 2025) and (Landsman et al., 2025)). Following the methodology of Amiram et al. (2022), the benchmark is constructed as the average log TVL over the pre-event estimation window $[-35, -16]$, which is sufficiently distant from the event to avoid contamination by anticipation effects. This normalization removes platform-specific time-invariant effects across protocols. Formally, for protocol i, the daily abnormal TVL (*AbTVL*) is defined as:

$$\text{AbTVL}_{i,t} = \log(\text{TVL}_{i,t}) - \frac{1}{20} \sum_{s=-35}^{-16} \log(\text{TVL}_{i,s})$$

We then aggregate daily abnormal TVL over various symmetric event windows to obtain the Cumulative Abnormal TVL (*CATVL*). For an event window $[t_1, t_2]$, this is given by:

$$\text{CATVL}_{i,[t_1,t_2]} = \sum_{t=t_1}^{t_2} \text{AbTVL}_{i,t}$$

We use a sequence of symmetric short-term event windows, ranging from [0] to $[-7, +7]$, to capture both immediate and delayed market reactions to open-source code disclosures, while accounting for potential information leakage leading up to the event. To evalute statistical significance, we conduct parametric t-tests across protocols. Although the raw distribution of TVL is highly skewed, log-transforming the data and aggregating over event windows improves normality and reduces the influence of outliers,

justifying the use of parametric inference procedures.

As a robustness check, we implement a market-model adjustment procedure to estimate expected TVL movements around repository launch events. This approach enables us to isolate the protocol-specific abnormal component of TVL changes associated with open-source disclosure, net of broader market fluctuations. Specifically, we estimate the following platform-specific model over the 75-day estimation window $[-90, -16]$ for each protocol i:

$$\log(\mathrm{TVL}_{i,t}) = \alpha_i + \beta_i \cdot \log(\mathrm{TVL}_t^{\mathrm{market}}) + \varepsilon_{i,t}$$

where $\mathrm{TVL}_t^{\mathrm{market}}$ denotes the log of Ethereum ecosystem's aggregate TVL on day t. Using the estimated coefficients $\hat{\alpha}_i$ and $\hat{\beta}_i$, we compute the daily *expected TVL* as:

$$\log(\widehat{\mathrm{TVL}}_{i,t}) = \hat{\alpha}_i + \hat{\beta}_i \cdot \log(\mathrm{TVL}_t^{\mathrm{market}})$$

The *market-adjusted abnormal TVL* is then defined as:

$$\mathrm{AbTVL}_{i,t}^{\mathrm{MktAdj}} = \log(\mathrm{TVL}_{i,t}) - \log(\widehat{\mathrm{TVL}}_{i,t})$$

Finally, we compute cumulative abnormal TVL based on these market-adjusted values across a sequence of symmetric event windows, ensuring that our estimates are not confounded by aggregate market-wide movements.

2.3.2 Panel Analysis

The above event-based design enables us to isolate the market and user behavioral responses to the open-source disclosure event, while mitigating confounding factors and ensuring temporal alignment across observations. Next, we turn to panel regression analysis, which allows us to exploit the full panel structure of our data and control for a richer set of observable and unobservable characteristics. We estimate the following OLS specifications:

$$Y_{i,t} = \alpha_i + \beta \cdot \mathrm{POST}(\textsc{OpenSourceLaunch})_{i,t} + \mathbf{X}_{i,t}'\Gamma + \gamma_t + \theta_i + \epsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ denotes the abnormal TVL(log) of platform i on day t, defined as the difference between the logarithm of the actual TVL and the average log TVL during the 20-day pre-event window $[-35, -16]$. The key explanatory variable, $POST(OpenSourceLaunch)_{i,t}$, is an indicator equal to one for post-event observations ($t \geq 0$), and zero otherwise. The vector $\mathbf{X}_{i,t}$ includes time-varying control variables as described in Section 2.

It is important to note that open-sourcing is not randomly assigned; the timing is a strategic choice by the startup. To mitigate concerns that $POST(OpenSourceLaunch)$ might capture other concurrent changes or that more successful startups choose to open-source (reverse causality), we rely on the high granularity of our data and fixed effects structure. The inclusion of industry fixed effects (capturing time-invariant differences across industry, such as business model or typical technology feature), α_i, means that identification comes from within-industry changes, so any static differences between open-source-friendly industry and others are differentiated out. Time fixed effects (capturing shocks common to all startups on day t, such as overall market sentiment), γ_t, ensure that economy-wide factors (like a bull market in crypto or a regulatory announcement affecting all DeFi projects) are accounted for. Blockchain fixed effects (capturing common product characteristics from specific blockchain infrastructures, such as protocol composability, and fee regimes), θ_i, ensure that persistent heterogeneity associated with the technological and economic environment of different blockchains is held constant. Thus, β captures the average change in *Abnormal TVL* following the open-source disclosure, comparing each startup's performance before and after the repository launch date.

As a robustness check, we estimate alternative specifications that include repository-level and time fixed effects. Repository fixed effects absorb all time-invariant characteristics specific to each repository, including its affiliated protocol, code architecture, developer team, and baseline visibility. Since each protocol's identity is embedded in its repositories, this approach also effectively controls for unobserved protocol-level heterogeneity. By controlling for repository-level fixed effects, β is identified from within-repository temporal variation. This removes the influence of time-invariant repository-specific characteristics, allowing for a cleaner estimation of the causal effect of open-source code disclosure on user engagement and platform activity. We estimate this model over various symmetric event windows (e.g., $[-3, +3]$, $[-5, +5]$, $[-7, +7]$, and $[-14, +14]$) to assess the robustness of the estimated treatment effect. Standard errors are clustered at the repository level to account for within-platform correlation.

3 Analysis and results

3.1 Event Study

Table 2 quantifies the real effects of open source across various event windows. We report the mean value of the natural logarithm of cumulative abnormal TVL(CATVL) for short-run windows from 0 to 7 days around the new repository launching date, along with t-statistics. We observe a consistent and statistically significant positive CATVL across all event windows. For example, the average cumulative abnormal TVL (log) within the [-1, +1] window is 0.136 (t = 4.47), increasing to 0.675 (t = 4.50) over the [-7,+7] window (= $ 4.817 million U.S. dollars). This persistent upward trend suggests that open-source disclosures foster user engagement as it signals DeFi protocol commitment to expanding services. The magnitude of responses is considerably larger when we restrict the sample to the first repository launch events. The [-1,+1] window exhibits a mean response of 0.732 (t = 2.67), and this grows to 4.064 (t = 3.02) in the [-7,+7] window (= $ 245.814 million U.S. dollars). This pronounced effect likely reflects the greater informational value and surprise associated with a protocol's initial transition toward open-source development. For many users, the first repository serves as a public commitment to code transparency, development credibility, and service expansion, especially in a DeFi environment characterized by information asymmetries and audit risk. Panel B shows the robustness results using a market-adjusted abnormal TVL measure. While statistical significance slightly attenuates in some windows, the core pattern remains: abnormal TVL increases significantly following repository launches.

Overall, these results provide strong support for the hypothesis that open-sourcing code enhances user engagement by reducing informational frictions and improving perceived protocol reliability. The consistent, statistically significant increases in CATVL across multiple event windows suggest that users respond favorably to transparent development practices, particularly when those disclosures represent a shift toward open collaboration. The magnified responses to first-time repository launches—both in economic magnitude and statistical strength—highlights the signaling value of initial code transparency. In environments like DeFi, where platform reliability, security, and developer intent are difficult to verify ex ante, such disclosures serve as credible commitments that lower user uncertainty and promote adoption. Notably, the robustness of these findings under market-adjusted benchmarks confirms that the observed effects are not merely reflections of aggregate market dynamics. Instead, they reflect platform-specific im-

provements in perceived legitimacy and service potential. Taken together, these findings highlight that open-source disclosures function not only as technical milestones, but also as trust-building mechanisms that can materially shape platform growth trajectories in decentralized ecosystems.

[Table 2 here]

We perform a series of robustness checks to validate the integrity of our event study results. Specifically, to address concerns that the observed effects might be confounded by simultaneous events, we exclude all repository disclosure events that coincide with other major announcements for the same protocol–such as high-profile partnership deals, token listings on major exchanges, or publicly reported security incidents (e.g., hacks or bug disclosures). These concurrent events could potentially attract investor attention and influence user behavior independent of code transparency. After excluding such overlapping events, we re-estimate the average treatment effect and find that the pattern of abnormal TVL remains consistent and statistically significant.

3.2 Panel Regression

Table 3 presents estimates from the panel regressions as specified in Equation 1. The dependent variable is the log abnormal TVL, computed as the deviation of actual TVL from its benchmark level based on pre-event performance. We estimate models over a sequence of symmetric event windows, ranging from short windows such as $[-3, +3]$ to broader windows like $[-14, +14]$ days surrounding the repository launch. Columns (1), (3), (5), and (7) report results from specifications that include repository-level fixed effects. Columns (2), (4), (6), and (8) instead include industry and blockchain fixed effects. Calendar month fixed effects are included in all columns to absorb macro-level shocks. Across all specifications, the coefficients on the indicator variable *POST (New repository)* are positive and statistically significant at the 5% level, suggesting that open-sourcing is associated with a significant increase in user commitment. In terms of economic magnitude, a typical DeFi protocol experiences a 1.6% increase in Total Value Locked (TVL) in the two weeks following the launch of a new open-source repository, relative to the two weeks preceding the launch. Moreover, the gradual increase in estimated coefficients as the event window expands suggests that user adoption continues to build over time, reflecting a persistent response to open-source disclosure.

For the estimated coefficients of control variables, the number of blockchains (log of chains) is positively associated with abnormal TVL and significant, suggesting that multi-chain deployment enhances user appeal and platform reach. The coefficient of DAO is 0.058, suggesting that platforms governed by a decentralized autonomous organization exhibit a 5.8% higher abnormal TVL. The presence of venture capital backing is significantly positive across all models, indicating that VC-backed protocols experience roughly 11% greater abnormal TVL, possibly reflecting investor confidence or stronger project marketing. In contrast, audit status and exchange listing show significant negative associations. The coefficient on *AUDIT* is negative and significant at the 10% level in all columns, indicating that more vulnerable or risk-conscious platforms are more likely to undergo audits but still face user skepticism. The negative and significant coefficient on *LISTED* suggests that platforms with existing tokens on exchanges may attract less incremental user lock-up following open-source disclosure, potentially due to investor attention being diverted toward speculative trading rather than platform use. Overall, these panel estimates support the event study findings and demonstrate a positive and persistent effect of open-source code publication on platform engagement as measured by user-committed capital.

4 Addressing Endogeneity

A major empirical challenge in identifying the causal effect of open-source code disclosure on user adoption in DeFi protocols stems from the endogeneity of disclosure decisions. Open-sourcing is voluntary and potentially strategic. Two primary sources of endogeneity are particularly salient: omitted variable bias and reverse causality. First, omitted variable bias arises when unobserved factors jointly influence both the likelihood of open-sourcing and the level of user adoption. For instance, protocols with more experienced development teams, stronger reputations, or superior technology stacks may be both more willing to release their source code and more attractive to users–independent of disclosure itself. If these latent characteristics are not fully observable or controlled for, the estimated effect of open-sourcing on platform performance may be biased upward. Second, reverse causality presents another concern. Rather than disclosure leading to increased user engagement, protocols might choose to open-source in response to rising TVL or user interest. For example, a platform experiencing rapid growth may open-source to signal technical quality, attract developers, or sustain momentum. Conversely, during market downturns or periods of uncertainty, teams may delay open-sourcing to avoid exposing potential weaknesses. In both

cases, causality runs from user behavior to disclosure rather than the other way around. To mitigate these endogeneity concerns, we implement three complementary empirical strategies.

4.1 Difference-in-Differences Analysis: Close-to-Open Transitions

Our first identification strategy exploits a sharp and well-defined event: the transition from closed-source to open-source status among a subset of DeFi protocols.[15] This transition allows us to isolate the causal effect of initiating code disclosure on user engagement within a narrow symmetric window around the disclosure event. The core identification idea is to compare changes in user engagement–proxied by abnormal Total Value Locked (TVL)before and after the transition from closed-source to open-source, relative to a matched control group of protocols that remained closed-source throughout the sample period. To construct a comparable counterfactual, we implement 1:1 nearest-neighbor propensity score matching based on a range of pre-treatment protocol characteristics, including platform size, multichain deployment, Twitter followers, oracle integration, DAO governance, venture backing, staking adoption, audit status, and exchange listing. The treated and control groups exhibit no statistically significant differences across these covariates after matching, as shown in Table A3. Each control protocol is assigned a pseudo event date corresponding to the actual open-source launch date of its matched treated protocol. We then examine user engagement over a symmetric event window–14 days before and after the actual or pseudo launch date–minimizing the influence of slow-moving confounders and reducing the risk of reverse causality. This difference-in-differences (DiD) design leverages the discrete timing of the initial repository disclosure and controls for time-invariant differences between protocols through matching and fixed effects. We estimate the following specification:

$$\log(\text{TVL}_{it}) = \alpha_i + \beta \cdot \textsc{Post}_{it} \times \textsc{Treated}_i + \mathbf{X}_{it}'\Gamma + \gamma_t + \theta_i + \varepsilon_{it}, \tag{2}$$

where Y_{it} is the log abnormal TVL for protocol i on day t, Treated is a dummy for protocols that eventually open-sourced, Post indicates days on or after the launch (or pseudo-launch) date, and the interaction term captures the average treatment effect of open-source transition. The vector $\mathbf{X}_i$ includes protocol-level covariates, while α_i, θ_i, and γ_t denote industry, blockchain, and year-month fixed effects,

[15] While the decision to open-source is ultimately endogenous, the discrete timing of a protocol's first public GitHub repository launch provides a tractable setting to examine short-run user responses.

respectively.

As shown in Table 4, the estimated coefficient on the interaction term *Treated (OpenSource)* × *POST (New repository)* is positive and statistically significant at the 10% level across both specifications, indicating that transitioning to open-source leads to a significant increase in user adoption. Column (1) includes protocol fixed effects, absorbing all time-invariant heterogeneity at the protocol level–such as differences in code architecture, team composition, or long-run strategic orientation. In terms of economic magnitude, treated protocols that transition to open-source experience a 16.2% increase in abnormal TVL in the 14 days following the launch of their first public GitHub repository, relative to their matched closed-source counterparts. Column (2), which instead includes industry and blockchain fixed effects, yields a nearly identical estimate (16.0%). These findings provide compelling evidence that code transparency serves as a credible signal of platform quality that encourages greater user commitment.

[Table 4 here]

To further address the possibility of reverse causality, we estimate a dynamic event-time specification centered around the repository launch date. Figure 2 plots the estimated treatment effects on abnormal log TVL for each day from $t = -10$ to $t = +10$, using day $t = 0$ (the launch date) as the reference point. The dynamic event-study estimates reveal no evidence of statistically significant differences in abnormal TVL between treated and control protocols during the pre-disclosure period, supporting the parallel trends assumption. Notably, we observe a marked and sustained increase in abnormal TVL beginning shortly after day 0, with the treatment effect stabilizing around +0.15 log points within approximately one week following the disclosure. The tight confidence intervals and consistent positive estimates across post-event days support a persistent treatment effect, rather than a transient or noisy spike. The absence of any significant pre-trends further rules out the possibility that treated protocols were already on a divergent growth trajectory prior to open-sourcing.

[Figure 2 here]

4.2 Instrumental Variables: GitHub Copilot AI Tooling Releases

Despite the fact that the transition from closed-source to open-source status partially mitigates concerns of reverse causality—by focusing on a discrete and time-stamped shift in disclosure status—this

transition remains an endogenous corporate decision. To further address the possibility of omitted variable bias that may confound our baseline estimates, we introduce an external instrumental variable (IV) strategy based on exogenous technological shocks originating from the GitHub platform. These releases represent a series of plausibly exogenous technological shocks to the broader open-source development environment, uncorrelated with DeFi market fundamentals.

GitHub Copilot was launched jointly by GitHub and OpenAI in June 2021 as the first large-scale AI-based pair programming assistant.[16] Directly integrated into widely used development environments (IDEs), Copilot leverages OpenAI's Codex model to provide context-aware code completions, assist with debugging, and simplify the development process. It has been rapidly adopted throughout the software development community. According to GitHub(2024), the tool now writes up to 46% of code in Copilot-enabled environments and increases developer productivity by approximately 55%.[17] These productivity gains substantially reduce the time, cost, and cognitive burden associated with writing, auditing, and refactoring code–benefits that are particularly pronounced for teams operating with limited engineering resources or heightened concerns over code quality and security.

Specifically, we identify five major release events of GitHub's Copilot-related AI development tools as plausibly exogenous technological shocks that lower the frictions of open-sourcing for DeFi projects. These milestones represent significant advances in AI-assisted coding, with broad implications for software engineering productivity, especially among resource-constrained teams. The five release events are detailed below:

1. **Copilot Technical Preview (June 29, 2021):** GitHub, in partnership with OpenAI, launched the technical preview of GitHub Copilot for Visual Studio Code.[18] This marked the first publicly available version of Copilot, enabling developers to experience real-time AI-driven code suggestions in popular IDEs. The launch sparked immediate interest and feedback from the open-source community.

2. **Copilot General Availability (June 21, 2022):** GitHub released Copilot for general commercial use, making it broadly available to individual developers for $10/month or to businesses via GitHub

[16] As of 2024, GitHub Copilot has been activated by over one million developers and adopted by more than 20,000 organizations. It has generated over three billion accepted lines of code, making it the world's most widely adopted AI-powered development tool.

[17] GitHub Blog: The Economic Impact of the AI-Powered Developer Lifecycle.

[18] GitHub Blog: Introducing GitHub Copilot – your AI pair programmer.

for Business subscriptions.[19] This release substantially expanded the user base and removed previous access barriers, accelerating adoption across startups and enterprises.

3. **Codespaces Free Tier (November 2022):** GitHub introduced a free tier for Codespaces, a cloud-based development environment that integrates directly with Copilot.[20] The free tier enabled developers to spin up pre-configured, cloud-hosted dev environments with Copilot enabled by default, drastically lowering infrastructure costs for testing and shipping open-source code.

4. **Copilot X Preview (March 2023):** GitHub announced the next-generation Copilot X, which incorporated GPT-4, a chat-based interface, documentation summarization, and code walkthroughs.[21] This preview expanded Copilot's utility beyond code suggestions, enhancing confidence in code correctness and simplifying debugging workflows-key concerns in open-source disclosures.

5. **Copilot Chat Release (December 2023):** Copilot Chat, a GPT-4-based conversational assistant embedded in IDEs, was officially released to the public.[22] The release further democratized AI access by offering conversational code assistance, making repository development and maintenance more accessible to small or non-technical teams.

To construct the instrumental variable, we define a monthly indicator variable, $\textsc{CopilotShock}_{it}$, which equals one if protocol i is observed in the same month or a month immediately following any of the identified GitHub Copilot release events. This window captures the period during which developers are most likely to adopt the new AI tooling and initiate new repository creation, reflecting the short-run behavioral response to external technological advancements. For this instrument to be valid, two conditions must be satisfied. First, in terms of relevance, the introduction of GitHub Copilot tools constitutes a major leap in AI-assisted programming. These tools substantially lower the cost, complexity, and effort associated with writing, auditing, and debugging code–especially smart contracts in decentralized applications. Such productivity gains are particularly valuable for smaller or resource-constrained DeFi teams, which may otherwise hesitate to disclose code due to quality concerns, cybersecurity risks, or limited engineering capacity. By enhancing developers' technical confidence and reducing disclosure frictions,

[19] GitHub Blog: GitHub Copilot is generally available to all developers.
[20] GitHub Blog: GitHub Universe 2022.
[21] GitHub Blog: GitHub Copilot X – the AI-powered developer experience.
[22] GitHub Blog: GitHub Copilot Chat now available in GA.

these tools plausibly raise the propensity to open-source. As illustrated in Figure 1, we observe a noticeable increase in the number of new repositories launched in the months immediately following major Copilot-related releases, suggesting that these technological shocks are associated with heightened code disclosure activity among DeFi protocols.

Second, regarding exogeneity, the timing and content of Copilot releases are determined independently by GitHub and OpenAI and reflect general-purpose technological innovation in software engineering. The rollout of these tools was neither targeted toward blockchain applications nor influenced by any DeFi-specific market conditions. Moreover, Copilot adoption does not directly impact DeFi user behavior or protocol-level TVL unless it first affects a team's decision to disclose its code. As shown in table 7, the coefficient of our instrument variable is statistically insignificant across both models, supporting the validity of the exclusion restriction. Hence, the exclusion restriction is likely to hold: Copilot shocks influence DeFi platform outcomes only indirectly, by changing the likelihood of repository creation. Together, these properties make the Copilot release events a credible and externally induced source of variation for identifying the causal impact of open-source disclosure on user adoption.

To estimate the causal effect of open-source repository disclosure on DeFi protocol adoption, we implement a two-stage least squares (2SLS) regression using the instrument $\textsc{CopilotShock}_{it}$.The empirical specification consists of the following two stages:

$$\textsc{NewRepository}_{it} = \pi_0 + \pi_1 \cdot \textsc{CopilotShock}_{it} + \mathbf{X}'_{it}\Pi + \theta_{s(i)} + \theta_{b(i)} + \nu_{it} \tag{3}$$

$$\log(\text{TVL}_{it}) = \alpha + \beta \cdot \widehat{\textsc{NewRepository}}_{it} + \mathbf{X}'_{it}\Gamma + \theta_{s(i)} + \theta_{b(i)} + \epsilon_{it} \tag{4}$$

where $\textsc{NewRepository}_{it}$ indicates whether protocol i launched a new GitHub repository in month t, $\textsc{CopilotShock}_{it}$ is a binary indicator equal to one if a given protocol-month falls within the same or subsequent month of any Copilot-related AI tooling release, and $\log(\text{TVL}_{it})$ denotes the logarithm of Total Value Locked. $\mathbf{X}_{it}$ includes protocol-level control variables. $\theta_{s(i)}$ and $\theta_{b(i)}$ denote industry and blockchain fixed effects, respectively. Standard errors are clustered at the protocol level.

[Table 5 here]

Table 5 presents the results. Column (1) shows the first-stage regression: the coefficient on CopilotShock is positive and statistically significant at the 1% level, with a first-stage F-statistic of 10.4, exceeding

conventional thresholds for weak instrument concerns.[23] This confirms that the release of Copilot-related tools significantly increases the likelihood of new open-source repository disclosures. Columns (2) and (3) report the second-stage regressions, with current and one-month-ahead log abnormal TVL as dependent variables. In the second stage, instrumented open-source disclosure is associated with a statistically and economically significant increase in user adoption. These results reinforce the interpretation that open-source code disclosure causally boosts user engagement, and validate the GitHub Copilot AI releases as a credible and relevant instrument.

4.3 Placebo Tests

Another concern in our empirical design is that open-source launches may coincide with other confounding factors that simultaneously affect user adoption, such as industry-specific shocks, token price cycles, or macro-level events that influence DeFi activity. For instance, a surge in market-wide enthusiasm for decentralized lending protocols may induce simultaneous increases in both repository creation and TVL, making it difficult to attribute the observed effects solely to open-sourcing. To address this concern, we implement a placebo test that follows standard practice in event studies. Specifically, we randomly reassign pseudo launch dates to each repository from the empirical distribution of actual launch dates. This preserves the seasonality and time distribution of open-source activity but breaks the direct link between the true event and observed outcomes. Using these placebo event dates, we recalculate abnormal TVL(log) within the $[-14, +14]$ event window and estimate the same regression as in Table3. We collect the estimated coefficient and t-statistic on the post-event indicator from each placebo regression. As shown in Figure 3, the distributions of the placebo coefficients and t-statistics are centered around negative values, rather than zero. This asymmetry reflects the underlying time trend in the DeFi ecosystem. As illustrated in Figure A1, the aggregate Total Value Locked (TVL) across protocols peaked in 2021 and exhibited a sharp and sustained decline thereafter. As a result, randomly assigned launch dates-which often fall in the post-2021 contraction period-tend to be associated with negative abnormal TVL movements. Moreover, the actual coefficient from the true specification, illustrated by the red dashed vertical

[23] The effect is broad-based: roughly 38% of protocols in our sample had active GitHub repositories at the time of a Copilot update, and over 60% of these published new code within six months. We find no evidence that Copilot releases directly affected TVL in the absence of disclosure, supporting the exclusion restriction.

line, lies far in the right tail of the simulated distribution, indicating that the observed effect is highly unlikely to have arisen by chance. This exercise supports the interpretation that the increase in cumulative abnormal TVL is causally linked to the timing of open-source repository disclosure, rather than driven by coincidental shocks.

[Figure 3 here]

5 Disclosure Costs and Long-Run Implications

5.1 Mimicking Risks

A fundamental tradeoff in open-source disclosure lies between signaling quality and exposing proprietary knowledge to potential imitators. This tension is particularly salient in decentralized finance (DeFi), where permissionless architectures and modular smart contracts allow new entrants to rapidly fork and replicate successful protocols. From a Schumpeterian perspective, the competitive advantage of innovation is inherently temporary: once disclosed, innovative rents are eroded as rivals imitate and improve upon the original design (Schumpeter, 1942).

To empirically examine this tension, we study a natural experiment stemming from the expiration of Uniswap V3's Business Source License (BSL) in April 2023.[24] Prior to this date, Uniswap V3's source code was legally protected against commercial reuse. Once the license expired, a wave of copycat protocols emerged, replicating Uniswap's liquidity management innovations. This shock offers a quasi-experimental setting to quantify the costs of mandatory disclosure in an environment of low IP protection.

[Figure 4 here]

Figure 4 illustrates the dramatic impact on industry structure. Panel (a) shows a sharp and sustained drop in the Herfindahl-Hirschman Index (HHI) of the decentralized exchange (DEX) industry immediately after the BSL expiration. Market concentration fell by over 50% within one month, indicating rapid erosion of Uniswap's dominant position. Panel (b) confirms a corresponding surge in the number of

[24] See details in Appendix B.

new DEX protocols launched post-shock, consistent with opportunistic market entry enabled by code copying.

We formally test the mimicking cost using a difference-in-differences framework. Specifically, we define a six-month symmetric event window surrounding the Uniswap V3 license expiration (October 2022 to October 2023). The sample includes all DeFi protocols that launched a new GitHub repository within this window and operate on at least two major blockchains. Protocols are assigned to the treated group if they belong to the DEX industry, and to the control group otherwise (e.g., lending, yield, derivatives, insurance). We compare the performance of new GitHub repositories launched by DEX protocols (treatment group) and non-DEX protocols (control group) in the six months before and after April 2023. We evaluate two outcome dimensions: (i) user adoption, measured by cumulative abnormal Total Value Locked (TVL) over windows of $[-1, +1]$, $[-7, +7]$, and $[-14, +14]$ days around the repository launch; and (ii) market investor response, measured by cumulative abnormal token returns over the same windows, benchmarked against a DeFi-wide token index.

[Table 6 here]

Table 6 reports the results. Panel A shows that new repositories by DEX protocols experience significantly larger cumulative abnormal TVL relative to the control group, suggesting that open-source imitation in competitive sectors like DEX can meaningfully shift user engagement and liquidity flows. Panel B reveals a similar pattern in token returns: post-Uniswap DEX repositories exhibit significantly higher market-adjusted returns, suggesting investor excitement and capital reallocation toward these forked platforms. These findings highlight a downside of transparency: when code is disclosed in a highly replicable form, it may catalyze aggressive entry and rent dissipation. The key assumption for DID validity is parallel trends–that in the absence of the Uniswap V3 license expiration, the evolution of abnormal TVL would have been similar between DEX and non-DEX protocols. As shown in Figure 5, we provide empirical evidence supporting the parallel trends assumption during the pre-treatment period. While open-sourcing is often lauded for promoting trust and community engagement, our results reveal that it can also expose firms to imitation threats that erode first-mover advantages. In line with Schumpeterian theory, the process of "creative destruction" is accelerated when knowledge spillovers occur without sufficient appropriation mechanisms.

[Figure 5 here]

5.2 Cybersecurity Risks

Next, we analyze the effect of open-source disclosure on DeFi protocol cybersecurity, specifically the likelihood and magnitude of subsequent protocol-level hacking incidents. This analysis is motivated by the growing tension between transparency and vulnerability in the Web3 ecosystem: while open-sourcing may facilitate community contributions and user adoption, it may also expose potential attack vectors to hackers. Figure 6 presents the quarterly distribution of total DeFi hack losses between 2020 and 2024. We observe several prominent spikes in hacking activity, most notably in 2021Q3 and 2022Q1. The 2021Q3 peak corresponds to the infamous *Poly Network Hack*, in which approximately $610 million were stolen through a cross-chain bridge exploit. In 2022Q1, the *Ronin Bridge Hack* led to another $620 million loss, exploiting validator access vulnerabilities tied to the Axie Infinity ecosystem. These large-scale events underscore the systemic risk and severity of DeFi protocol hacks.

[Figure 6 here]

To assess whether the decision to launch a public GitHub repository is associated with greater cybersecurity risk, we conduct a cross-sectional comparison based on a matched sample. We manually collect a dataset of 360 DeFi protocol hack events that occurred between 2020 and June 2025. Details of the events are provided in Appendix C. The treated group consists of protocols that launched their first public repository within a given year, while the control group includes protocols that never disclosed their code. We employ 1:1 nearest-neighbor propensity score matching based on key pre-treatment characteristics, including protocol size, social media following, multichain deployment, governance structure, audit status, and venture capital backing. After matching, we obtain a balanced sample of 287 treated and 287 control protocols. We then compare the subsequent likelihood of being hacked across the two groups to evaluate whether open-source activity is associated with increased attack exposure.

[Table 7 here]

Table 7 presents the results. Column (1) estimates a linear regression model with *Hack Loss (log)* as the dependent variable, capturing the total dollar amount of funds lost to attacks (in log scale). Column (2) reports a logistic regression where the outcome is a dummy equal to one if the protocol experienced at least one hack during the post-treatment window. In both specifications, we include controls as well

as industry, blockchain, and year fixed effects. Standard errors are clustered at the protocol level. The coefficient on *Treated* is positive and significant across both specifications. Compared to protocols that never open-sourced their code, those that launched a public repository experienced an average increase of 76.8% in hack losses and a 4.6 percentage point higher likelihood of being hacked in the future. These results suggest that open-sourcing is associated with increased attack surface or visibility to hackers. Economically, this implies a potential trade-off: while transparency may boost credibility and adoption, it can also raise protocol-level risk exposure.

Beyond evaluating whether open-source protocols are more likely to be targeted by attackers, we also examine how users respond to such perceived security risks. Specifically, we compare user reactions to repository launches that occur during periods of heightened hacking concerns versus more stable times. To visualize these dynamics, Figure A2 plots the evolution of abnormal TVL (in logs) around open-source repository launch events, spanning a [-14, +14] day window. Panel (b) separates the sample based on exposure to major security breaches in the ecosystem at the time of the repository launch. The results reveal a striking contrast: protocols that launch repositories during high-profile hack periods experience a muted, or even negative, TVL response. In contrast, protocols launching during low-risk periods continue to see a steady increase in abnormal TVL. These patterns suggest that perceived cybersecurity threats can significantly dampen the positive signaling effect of code transparency, underscoring the conditional nature of user trust in open-source innovation.

5.3 Token Returns

While the previous section shows that open-sourcing a new repository leads to a significant increase in platform adoption—as evidenced by a rise in Total Value Locked (TVL)—we now turn to how investors respond to the same events through the lens of token prices. Whereas users appear to reward transparency and signaling, investors may perceive open-source disclosure differently, especially in the short term. To evaluate these reactions, we conduct an event study of cumulative abnormal returns (CARs) surrounding new repository launches.

[Table 8 here]

Table 8 presents results from the market-adjusted CAR analysis. Panel A focuses on short-term windows around the event date. We find that token prices drop significantly in the immediate aftermath of

open-source announcements. CARs are negative and statistically significant in the [0] window (−0.80%, $t = -3.70$), as well as in [−1, +1] (−0.96%, $t = -3.13$) and [−2, +2] (−1.59%, $t = -3.66$). In the [−3, +3] window, the return remains negative (−0.57%), though not statistically significant. These patterns suggest that some investors sell their tokens due to concerns about copycat competition or heightened cyber vulnerabilities, indicating that short-term traders are unwilling to finance open-source projects whose returns materialize only over longer horizons.

Panel B examines longer-term horizons and reveals a striking reversal. CARs turn significantly positive over all extended windows, increasing from 7.9% over [0, 90] days ($t = 1.84$), to 13.3% over [0, 120] days ($t = 2.17$), 14.7% over [0, 150] days ($t = 2.16$), and ultimately reaching 17.9% over [0, 180] days ($t = 2.33$). These effects are not only statistically significant but also economically meaningful, indicating that the market gradually revalues open-source disclosure as its benefits become evident through user growth, developer engagement, and platform performance, thereby aligning the investor base more closely with long-term value creation.

The evidence points to a clear long-term benefit: open-source code sharing is associated with greater platform adoption and ultimately higher valuation. Yet, we also uncover a short-term cost in the form of a negative immediate market reaction. These findings are consistent with frameworks of investor misvaluation and behavioral finance. Investors appear slow to process innovation-related signals embedded in code disclosures. As Cohen et al. (2013) show in public equity markets, firms with strong innovation track records are often initially underpriced, leading to alpha opportunities for informed investors. Our setting offers a real-time analog: investors underreact to open-source events, only to adjust their beliefs later when observable performance catches up.

This misvaluation is exacerbated by the phenomenon of "code-washing" documented by Gefen et al. (2024), where some startups strategically release meaningless or plagiarized code to appear innovative. In such an environment, investors face an adverse selection problem and may discount all open-source activity. Genuine projects thus do not receive full credit immediately, but are gradually recognized as their performance diverges from superficial imitators. The sustained long-term outperformance we document among open-sourcers fits this dynamic. From a practical perspective, this opens up an opportunity: sophisticated investors with the technical literacy to evaluate code quality and developer activity may be able to arbitrage this mispricing.

The evidence of short-term price declines following open-source events can feed into managerial

decision-making. If founders anticipate this reaction, those who are more concerned about short-run stock (or token) performance might delay open-sourcing or avoid it entirely. Our ancillary finding that startups with short-horizon investors delay code release is telling—these managers likely feel acute pressure not to rock the boat in the near term. The behavior aligns with the survey evidence of Graham et al. (2005), where executives openly acknowledged they would postpone or avoid valuable long-term projects if it meant missing their quarterly targets. In the startup world, there are analogous pressures: a founder might fear that a dip in token price could derail an upcoming token sale, or that it could be interpreted as a sign of weakness by competitors. Such fears, whether justified or not, create a bias toward the status quo (closed-source development) even when open-source could be advantageous in the long run. Our findings thus not only shed light on how investors interpret open-source events, but also contribute to a broader understanding of how financial markets process technological disclosures.

5.4 Product Innovation: Long-term Feature Creation

To complement our earlier analysis of short-term user responses to code disclosure, we next investigate whether open-source repository launches also contribute to long-term innovation outcomes. Specifically, we examine whether making a protocol's codebase publicly available fosters the development of new features over the months following disclosure. The economic rationale is that open-sourcing enhances transparency, invites external developer engagement, and increases the composability of protocol code–conditions that are conducive to innovation. By reducing the costs of third-party audits, encouraging community suggestions, and enabling modular reuse of code, open repositories may serve as catalysts for product expansion. We define new features as substantive additions to the protocol's functionality, such as the launch of staking or lending modules, cross-chain deployment, oracle integrations, or other technical upgrades that extend the protocol's scope or user experience. These changes not only improve the platform's utility but may also attract new users and liquidity, reflecting a deeper and more persistent growth channel beyond the initial disclosure event.

We formally test this long-term performance using a difference-in-differences (DID) framework. The treated group consists of protocols that launched their first open-source repository in a given year, while the control group includes protocols that never opened a repository throughout the sample period. Each treated protocol is matched 1:1 to a control protocol using nearest-neighbor propensity score matching, based on key protocol characteristics including protocol size, number of supported chains, social media

follower base, DAO structure, oracle, venture capital backing, staking scale, audit status, and listing status. To ensure comparability in time exposure, each control protocol is assigned the same event year as its matched treated counterpart. The analysis focuses on a [-2, +2] event window surrounding the first repository launch. The dependent variable is *New feature (log)*, defined as the log of the number of new features introduced by each protocol in a given year. *Post* is an indicator variable equal to one if the year is equal to or after the first GitHub repository launch date for a given protocol, and zero otherwise. Our coefficient of interest is the interaction term $Treated \times Post$, which captures the differential change in the number of new features developed after the first repository launch for treated protocols, relative to the control group that never launched a repository.

[Table 9 here]

Column (1) of Table 9 reports baseline DiD results, while column (2) implements a dynamic DiD analysis. The estimated coefficient on $Treated \times Post$ is 0.328 and statistically significant at the 1% level, suggesting an approximate 39% increase in the number of new features developed by treated protocols after open-source adoption, relative to the matched control group. In column (2), the coefficients on $Treated \times After^{1yr}$ and $Treated \times After^{2yr}$ are both positive and statistically significant, indicating a post-treatment increase in the number of new features. Importantly, the pre-trends (i.e., $Before^{1yr}$ and $Before^{2yr}$) are not statistically significant, supporting the validity of the parallel trends assumption. This pattern is consistent with the interpretation that open-sourcing promotes protocol improvement and functional diversification, thereby contributing to greater user adoption and increased total value locked (TVL).

6 Mechanisms

In this section, we explore the mechanisms through which open-sourcing may influence user adoption. First, the governance and monitoring channel posits that the institutional structure of a protocol shapes how effectively code transparency signals quality and long-term commitment. Second, the innovation signaling channel suggests that disclosing source code allows protocols to convey the strength, extensibility, and transparency of their underlying technology, thereby reducing information asymmetry and building user trust. Third, the developer engagement and network channel highlights how code disclosure fosters community contributions, accelerates iterative development, and expands the protocol's developer ecosystem.

6.1 Governance and Monitoring Channel

The effectiveness of code disclosure hinges not only on whether a protocol chooses to open-source its code, but also on how users interpret and respond to that disclosure based on the protocol's governance and oversight environment. In decentralized finance, where enforcement mechanisms are weak and platform quality is opaque, users face considerable uncertainty in capital allocation decisions. Effective governance serves as a crucial interpretive lens through which users assess the credibility and implications of code transparency. Protocols with strong governance enable stronger community monitoring and participatory oversight, which can amplify the credibility of code releases. In this way, governance acts as an interpretive lens through which users assess the salience and reliability of open-source disclosures.

To test this mechanism, we explore three organizational dimensions. First, we classify protocols by governance type, distinguishing between decentralized autonomous organizations (DAOs) and non-DAOs. DAOs typically feature token-based voting and participatory governance, enabling stronger community oversight and a higher demand for credible transparency. Second, we assess external certification by identifying protocols backed by venture capital (VC) investors. VC participation often signals professionalism and monitoring, amplifying the credibility of open-source disclosures. Third, we consider user risk exposure by focusing on lending protocols, where users entrust their assets to the platform for extended periods, making them more sensitive to code security and transparency. Next, we estimate the panel specifications (Equation 1) separately within subsamples defined by each organizational dimension. The dependent variable is the log of abnormal TVL over a $[-14, +14]$ day window surrounding the repository disclosure date. All regressions include protocol-level controls, as well as fixed effects for industry, blockchain, and year-month. To formally test for heterogeneity in the treatment effects across groups, we perform a Chow test comparing the subsample estimates.

[Table 10 here]

Table 10 presents the results. Panel A shows that protocols governed by DAOs experience a statistically significant increase in abnormal TVL following repository disclosure, whereas non-DAO protocols do not. This finding is consistent with the notion that decentralized governance strengthens community oversight and trust, which in turn enhances the credibility of code transparency and encourages greater user commitment and capital allocation. Panel B reveals that protocols backed by venture capital (VC)

investors exhibit a significantly stronger response to code disclosure compared to those without VC backing. VC involvement may signal professional oversight, due diligence, and a higher level of institutional quality, which enhances users' interpretation of the open-source event as a meaningful signal of security and legitimacy. Panel C shows similar heterogeneity based on user risk exposure. Compared with other industries, lending protocols–where users typically lock assets into smart contracts and face higher potential downside risk–exhibit a significantly larger impact of code disclosure on abnormal TVL. This pattern suggests that when users are more exposed to platform risk, they respond more strongly to transparency signals.

6.2 Innovation Signaling Channel

The effectiveness of code disclosure depends not only on whether platforms disclose, but also on the quality and nature of the innovation being revealed. In decentralized settings, where users and developers face substantial uncertainty, high-quality code can serve as a credible signal of a platform's technical competence, transparency, and long-term viability. Repositories with higher levels of developer popularity, greater extensibility, and stronger transparency enhance the salience of the disclosure, allowing ecosystem participants to more effectively evaluate platform value. In this way, code transparency plays a role akin to patents or R&D disclosures in traditional sectors, acting both as a commitment device and a signal of technological strength.

To examine how the quality of innovation affects the efficacy of code disclosure, we focus on three repository-level characteristics that proxy for distinct dimensions of innovation value: (1) Popularity, measured by the number of GitHub stargazers, which reflect community recognition and perceived importance. *HIGH Stargazers* refers to repositories in the top quartile of the star distribution, and *LOW Stargazers* includes the remaining repositories. (2) Extensibility, proxied by the number of forks, which indicates modularity and the potential for reuse (*HIGH Forks* = top quartile). (3) Transparency, proxied by whether the repository is *analytics-oriented*. A repository is defined as analytics-oriented, based on textual analysis of metadata fields, identifying references to data reporting, dashboard functionality, or user-facing metrics. Next, we estimate the panel specifications (Equation 1) separately within subsamples defined by each innovation quality dimension and perform a Chow test comparing the subsample estimates.

[Table 11 here]

Table 11 presents the results. Panel A shows that protocols whose repositories attracted high stargazer counts exhibit a significantly larger post-disclosure increase in abnormal TVL compared to their low-stargazer counterparts. Panel B reveals similar heterogeneity when grouping by the number of forks, reinforcing the idea that extensible codebases-those more frequently reused-amplify the benefits of transparency. Panel C further shows that repositories classified as analytics-oriented yield significantly greater abnormal TVL responses than non-analytics repositories. This supports the notion that interpretability and user-facing functionality enhance the salience and credibility of open-source disclosures. Together, these results affirm the innovation signaling channel. It is not simply the act of open-sourcing, but the visibility, reusability, and transparency of the disclosed innovation that drive user adoption and engagement.

6.3 Development Engagement and Network Channel

Beyond signaling innovation quality, open-source code disclosures can catalyze platform growth by fostering engagement within developer communities. In decentralized ecosystems, attracting and retaining contributors is crucial to sustaining innovation. GitHub repositories not only serve as disclosure vehicles but also as coordination hubs for collaborative activities such as pull requests, issue resolution, and code reviews. These community interactions may generate a virtuous cycle where transparency stimulates participation, which, in turn, boosts credibility and user adoption. Platforms embedded in denser developer networks—characterized by higher contributor activity, broader cross-platform engagement, or stronger collaboration ties—are more likely to benefit from this feedback loop.

To examine this mechanism, we categorize protocols based on three repository-level attributes that reflect development team engagement and network embeddedness: (1) Developer Team Size, measured by the total number of unique developers contributing to the repository. *High Size* repositories fall within the top quartile of this distribution, while *Low Size* repositories comprise the remainder. (2) Developer Productivity, proxied by the average number of commits per developer, with *High-Productivity* repositories in the top quartile. (3) Developer Network Connectivity, defined as the average number of distinct repositories to which each contributor has contributed across the platform. *High Network* repositories represent the top quartile of this metric. We estimate Equation 1 separately within each subgroup and

conduct Chow tests to formally assess whether treatment effects differ significantly across strata.

[Table 12 here]

Table 12 presents the results. Panel A shows that protocols with larger development teams benefit more from open-source launches. The estimated treatment effect for *High Size* protocols is over four times that of *Low Size* protocols, and the difference is statistically significant. This suggests that larger teams may offer superior pre-launch code quality, peer review, or post-launch support, which collectively enhance user trust. Panel B reveals even sharper heterogeneity based on developer productivity. For *High Productivity* teams—those with the most commits per contributor—the estimated effect of disclosure is 0.073, whereas the effect for *Low-Productivity* teams is statistically indistinguishable from zero. This finding implies that users are not merely reacting to code publication but are attuned to signals of developer expertise and reliability. Panel C indicates that protocols with highly connected developers—measured by contributors' average cross-platform activity—experience significantly stronger abnormal TVL gains following code disclosure. The effect for the *High Network* group is significantly positive, compared to a non-significant effect for the *Low-Network* group. This result underscores the role of network externalities: developers embedded in broader ecosystems likely enhance credibility, facilitate knowledge transfer, and draw attention from adjacent communities. Taken together, these results highlight that the benefits of code transparency depend on the developmental and network context. Simply releasing code is not sufficient. Rather, open-sourcing must be accompanied by signals of active engagement, team credibility, and integration within broader developer ecosystems.

7 Conclusion

This study examines open-source code disclosure as a strategic form of voluntary transparency in decentralized finance (DeFi). By releasing repositories on GitHub, startups go beyond coding to also signal commitment to transparency, foster user trust, and reduce information asymmetries. Using a novel panel dataset tracking hundreds of DeFi startups, their GitHub activity, and high-frequency economic performance, we show that open-source disclosures are associated with substantial and persistent growth in user engagement. Our identification strategy combines a matched difference-in-differences design with an instrumental variable approach based on the staggered rollout of GitHub Copilot, providing robust

causal evidence that code transparency leads to increased user adoption. Mechanism tests reveal that the effects are amplified for protocols with decentralized governance, venture capital backing, high-quality and extensible codebases, and engaged developer networks—highlighting the importance of governance credibility, innovation signaling, and network spillovers.

However, open-source disclosure also entails strategic costs. Transparent protocols face heightened exposure to security breaches and imitation: we document a higher likelihood of hacks post-disclosure and show that Uniswap's market share declined sharply—by over 50%—after its code license expired and competing clones emerged. Despite these risks, open-sourcing is positively associated with long-run innovation, as open protocols develop significantly more features and integrations in the years following disclosure. Taken together, our findings highlight a transparency dilemma: openness catalyzes growth and innovation but increases proprietary and cybersecurity risks. For entrepreneurs, this highlights the need to weigh short-term vulnerabilities against long-term strategic value; for investors, it suggests opportunities to identify underappreciated yet high-potential projects based on transparency signals. More broadly, we contribute to the voluntary disclosure and innovation valuation literature by quantifying the economic trade-offs of open-sourcing in highly innovative markets.

References

Aboody, D. and B. Lev (2000). Information asymmetry, r&d, and insider gains. *Journal of Finance 55*(6), 2747–2766.

Adams, H. (2020, September). Uniswap and the rise of the vampire attack. *Uniswap Blog*.

Allen, A., A. Lewis-Western, and K. Valentine (2022). Regulation and innovation: Evidence from the sarbanes-oxley act. *Journal of Accounting and Economics 73*(2-3), 101453.

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

Amiram, D., E. Lyandres, and D. Rabetti (2025). Trading volume manipulation and competition among centralized crypto exchanges. *Management Science 71*(3), 1369–1387.

Anton, J. J. and D. A. Yao (2002). The sale of ideas: Strategic disclosure, property rights, and contracting. *Review of Economic Studies 69*(3), 513–531.

Anton, J. J. and D. A. Yao (2004). Little patents and big secrets: Managing intellectual property. *RAND Journal of Economics 35*(1), 1–22.

Barth, M. E. and K. H. Gee (2024). Accounting and innovation: Paths forward for research. *Journal of Accounting and Economics 78*(1-2), 101733.

Beyhaghi, M., P. Khashabi, and A. Mohammadi (2023). Pre-grant patent disclosure and analyst forecast accuracy. *Management Science 69*(5), 3140–3155.

Bhattacharya, S. and J. R. Ritter (1983). Innovation and communication: Signalling with partial disclosure. *Review of Economic Studies 50*(2), 331–346.

Bhojraj, S., P. Hribar, M. Picconi, and J. McInnis (2009). Making sense of cents: An examination of firms that marginally miss or beat analyst forecasts. *Journal of Finance 64*(5), 2361–2388.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4), 1347–1393.

Botosan, C. A. and M. A. Plumlee (2002). A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research 40*(1), 21–40.

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

Campello, M., P. Jin, D. Rabetti, and F. Saleh (2023). The market for crypto zombies: Undercollateralization in DeFi lending. Working paper.

Cheng, C. S. A., P. Guo, C.-H. Weng, and Q. Wu (2021). Innovation and corporate tax planning: The distinct effects of patents and r&d. *Contemporary Accounting Research 38*(1), 621–653.

Cohen, L., K. Diether, and C. Malloy (2013). Misvaluing innovation. *Review of Financial Studies 26*(3), 635–666.

Coinmonks (2022, March). Total value locked (tvl) and why it matters in DeFi. *Medium*.

Cong, L. W., Y. Li, and N. Wang (2021). Tokenomics: Dynamic adoption and valuation. *Review of Financial Studies 34*(3), 1105–1155.

Cong, L. W., Y. Li, and N. Wang (2022). Token-based platform finance. *Journal of Financial Economics 144*(3), 972–991.

Cong, L. W., E. S. Prasad, and D. Rabetti (2025, April). Financial and informational integration through oracle networks. NBER Working Paper No. 33639.

Cong, L. W., D. Rabetti, C. C. Wang, and Y. Yan (2025, March). Centralized governance in decentralized organizations. Harvard Business School Working Paper.

Conti, A., C. Peukert, and M. Roche (2024). Beefing it up for your investor? engagement with open source communities, innovation and startup funding: Evidence from github. Unpublished manuscript, SSRN Working Paper ID 3883936.

Cream Finance (2020, August). Forked from compound: The launch of cream finance. *Medium*.

De Simone, L., P. Jin, and D. Rabetti (2024). Tax planning, illiquidity, and credit risks: Evidence from DeFi lending. Working paper, SSRN ID 4764605.

Dye, R. A. (1985). Disclosure of nonproprietary information. *Journal of Accounting Research 23*(1), 123–145.

Dyer, T. A., S. Glaeser, M. H. Lang, and C. Sprecher (2024). The effect of patent disclosure quality on innovation. *Journal of Accounting and Economics 77*(2-3), 101647.

Eberhart, A. C., W. F. Maxwell, and A. R. Siddique (2004). An examination of long-term abnormal stock returns and operating performance following r&d increases. *Journal of Finance 59*(2), 623–650.

Ellis, J. A., C. E. Fee, and S. E. Thomas (2012). Proprietary costs and the disclosure of information about customers. *Journal of Accounting Research 50*(3), 685–720.

Emery, L. P., C. Lim, and S. Ye (2024). The private value of open-source innovation. Working paper, SSRN ID 5054880.

Francis, J., R. LaFond, P. Olsson, and K. Schipper (2008). Costs of equity and earnings attributes. *Accounting Review 83*(4), 967–1010.

Gao, P., J. C. Wu, and J. L. Zimmerman (2017). Unintended consequences of sox: The effect on r&d spending. *Journal of Accounting and Economics 63*(2-3), 439–462.

Gefen, O., D. Rabetti, Y. Sun, and C. Zhang (2024). Code-washing: Evidence from open-source blockchain startups. Working paper, SSRN ID 5068292.

Glaeser, S. and W. R. Landsman (2021). Disclosure of trade secrets and innovation: Evidence from the dtsa. *Journal of Accounting and Economics 71*(2-3), 101365.

Glaeser, S. and M. Lang (2024). Measuring innovation and navigating its unique information issues: A review of the accounting literature on innovation. *Journal of Accounting and Economics 78*, 101720.

Glaeser, S., J. Michels, and R. E. Verrecchia (2020). Discretionary disclosure and manager horizon: Evidence from patenting. *Review of Accounting Studies 25*(2), 597–635.

Graham, J. R., C. R. Harvey, and S. Rajgopal (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics 40*(1-3), 3–73.

Harvey, C. R. and D. Rabetti (2024). International business and decentralized finance. *Journal of International Business Studies 55*, 840–863.

Healy, P. M. and K. G. Palepu (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics 31*(1-3), 405–440.

Heeley, M. B. and R. Jacobson (2008). The recency of technological inputs and financial performance. *Strategic Management Journal 29*(7), 723–744.

Heeley, M. B., S. F. Matusik, and N. Jain (2007). Innovation, appropriability, and the underpricing of initial public offerings. *Academy of Management Journal 50*(1), 209–225.

Hegde, D., K. Herkenhoff, and C. Zhu (2023). Patent publication and innovation. *Journal of Political Economy 131*(7), 1845–1903.

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

Kim, J. and K. Valentine (2021). Patent disclosure and innovation incentives: Evidence from the american inventor's protection act. *Journal of Accounting Research 59*(5), 1579–1624.

Landsman, W. R., E. Lyandres, E. L. Maydew, and D. Rabetti (2025, March). Auditing smart contracts. HKU Jockey Club Enterprise Sustainability Global Research Institute Archive.

Lee, S. W., J. Pinto, D. Rabetti, and G. Sadka (2024). Blockchain-induced supply chain transparency and firm performance: The role of capacity utilization. Working paper.

Lerner, J. and M. Schankerman (2010). The comingled code: Open source and economic complementarities. *Journal of Economics & Management Strategy 19*(2), 473–489.

Lerner, J. and M. Schankerman (2011, March). Managing the open source vs. proprietary decision. Harvard Business School Working Knowledge.

Lerner, J. and J. Tirole (2005). The economics of technology sharing: Open source and beyond. *Journal of Economic Perspectives 19*(2), 99–120.

Leuz, C. and P. D. Wysocki (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research 54*(2), 525–622.

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

Luo, M., D. Rabetti, and S. Yu (2024). Blockchain adoption and audit quality. Working paper.

Lyandres, E., B. Palazzo, and D. Rabetti (2022). Initial coin offering (ico) success and post-ico performance. *Management Science 68*(12), 8658–8679.

Merton, R. C. and R. T. Thakor (2023, February). Trust, transparency, and complexity. Harvard Law School Forum on Corporate Governance.

Messari (2023, August). How to interpret total value locked (tvl) in DeFi. Messari Research Report.

Nagle, F. (2019). Open source software and firm productivity. *Management Science 65*(3), 1191–1215.

Parlour, C. A. (2023). How safe is DeFi? systemic risk and fragility. Keynote Talk at the 2023 Global AI Finance Research Conference.

Plumlee, M., Y. Xie, M. Yan, and J. J. Yu (2015). Bank loan spread and private information: Pending approval patents. *Review of Accounting Studies 20*(2), 593–638.

Poly Network (2021, August). Poly network hack explained. *Poly Network Blog*.

Saggese, P., M. Fröwis, S. Kitzler, B. Haslhofer, and R. Auer (2025). Towards verifiability of total value locked (tvl) in decentralized finance. BIS Working Paper No. 1268.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. New York: Harper & Row.

Stablecoin Strategist (2023, October). How to read and analyze tvl data using defillama. *Dev.to*.

Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics 104*(4), 655–669.

Sygnum (2023, June). Total value locked (tvl): Measuring DeFi growth correctly. *Sygnum Blog*.

Taalbi, J. (2022). Innovation with and without patents. *arXiv preprint arXiv:2210.04102*.

Valentine, K., J. L. Zhang, and Y. Zheng (2025). Strategic scientific disclosure: Evidence from the leahy–smith america invents act. *Journal of Accounting Research*. Forthcoming.

Verrecchia, R. E. (1983). Discretionary disclosure. *Journal of Accounting and Economics 5*, 179–194.

Verrecchia, R. E. (2001). Essays on disclosure. *Journal of Accounting and Economics 32*(1-3), 97–180.

Figure 1. GitHub Copilot Tool Launches and New Repository Creation. This figure plots the monthly number of newly created GitHub repositories by DeFi protocols from 2021 to 2024. The vertical shaded areas mark the staggered release dates of major GitHub AI tooling products, including Copilot Preview (June 2021), Copilot General Availability (June 2022), Codespaces Free Tier (November 2022), Copilot X Preview (March 2023), and Copilot Chat Release (December 2023).

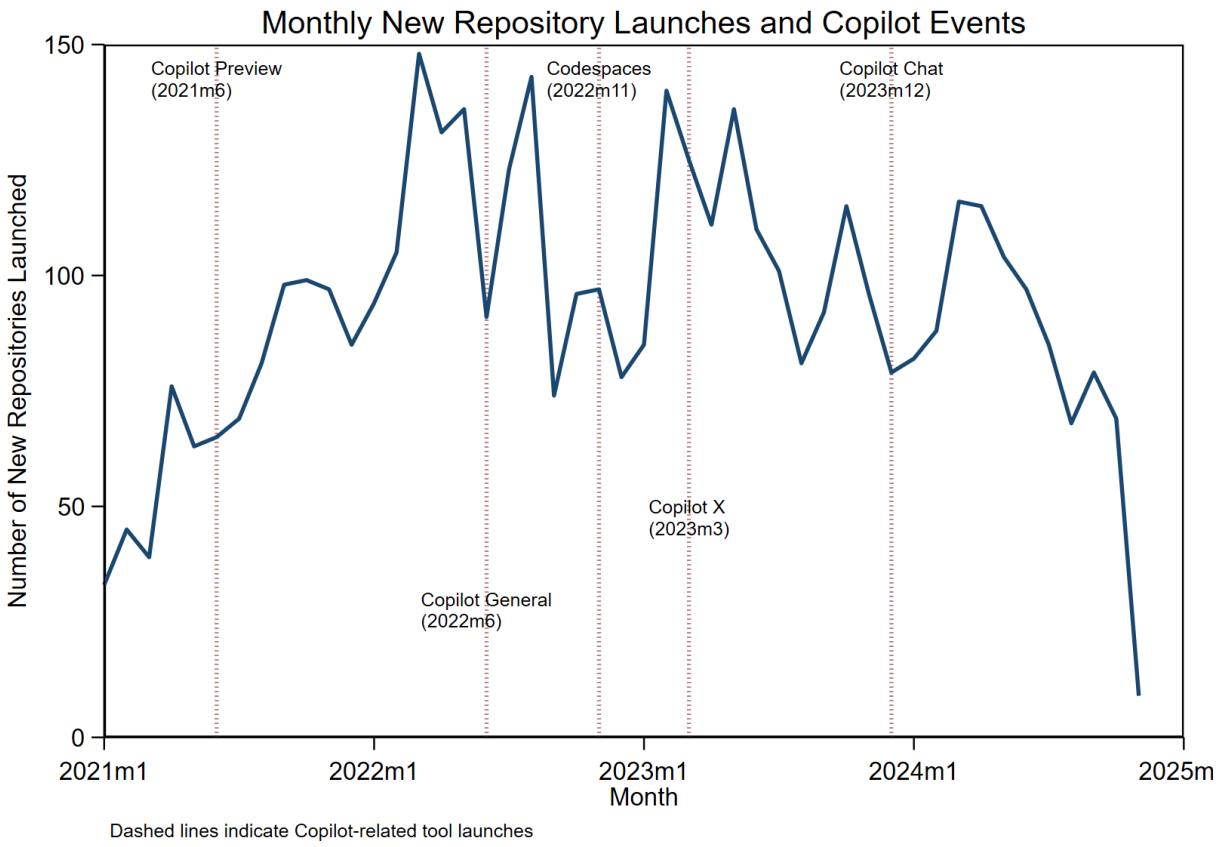


Figure 2. Dynamic Effects of Transition from Closed-Source to Open-Source on Abnormal TVL. This figure illustrates the dynamic treatment effects of transitioning from closed-source to open-source, measured by DeFi protocols' abnormal Total Value Locked (TVL) in logarithmic form. We estimate an event-time difference-in-differences specification, where the treated group consists of protocols that launched their first public GitHub repository, representing a shift from closed to open source, while the control group includes matched protocols that remained closed-source throughout the sample period. We implement 1:1 nearest-neighbor propensity score matching based on protocol characteristics including *SIZE*, *Chains(log)*, *Twitter followers(log)*, *ORACLE*, *DAO*, *Venture capital*, *Staking(log)*, *AUDIT*, and *LISTED*. Each control protocol is assigned a pseudo event date corresponding to the matched treated protocol's actual open-source launch, ensuring temporal comparability. The x-axis shows days relative to the open-source launch (day 0 as baseline). Each dot plots the estimated interaction coefficient between treatment status and event day. Vertical bars represent 90% confidence intervals. All regressions include protocol-level controls and fixed effects for blockchain, industry, and calendar month.

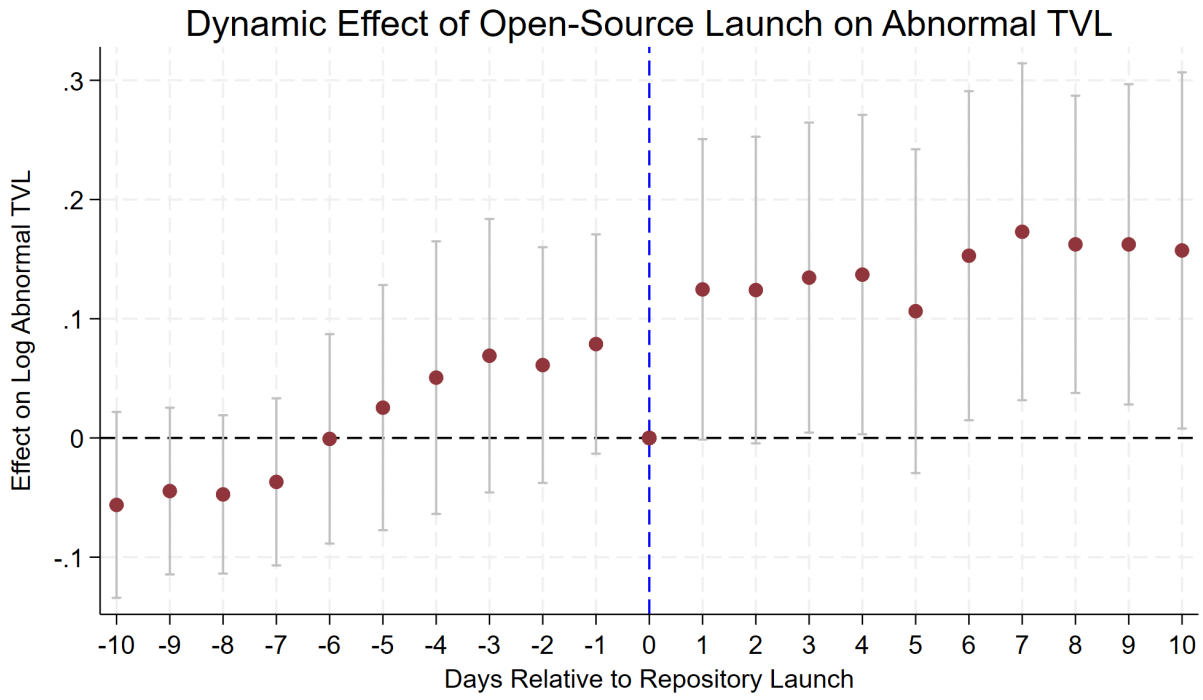
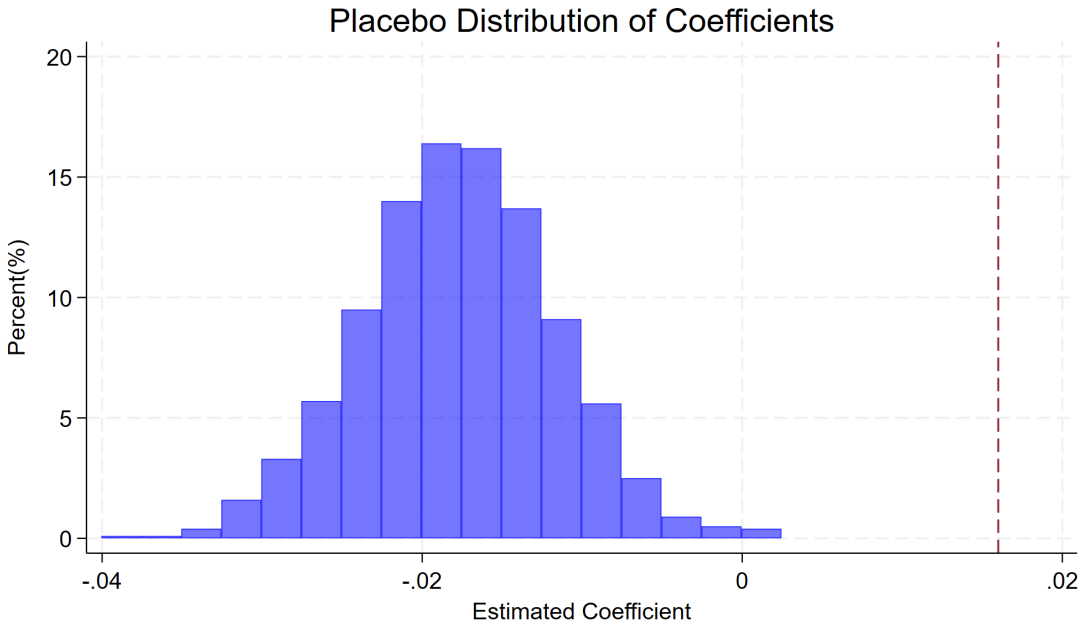
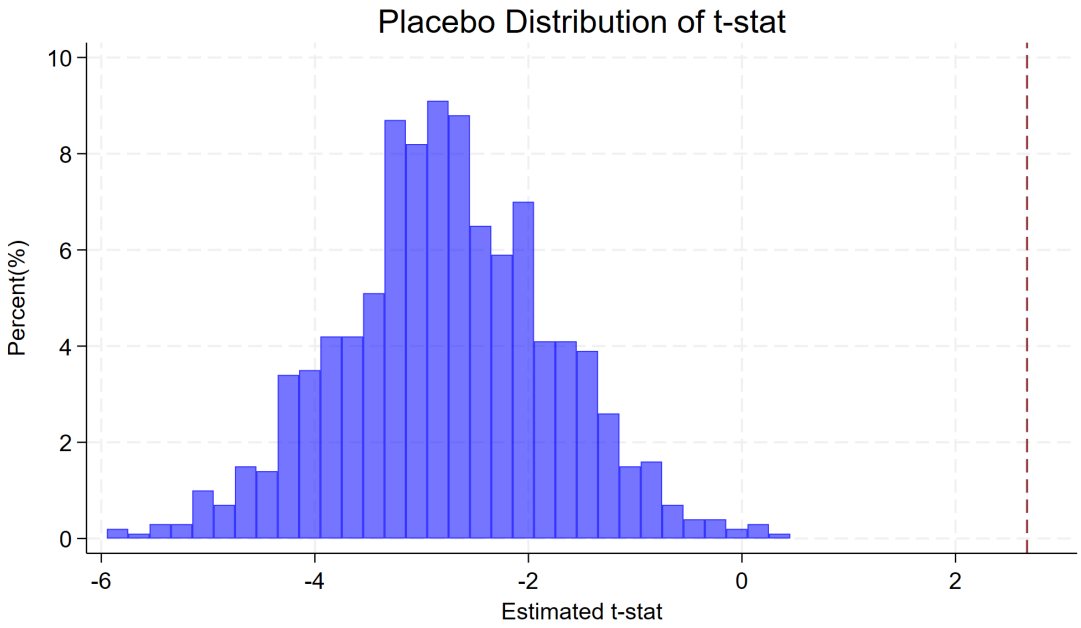


Figure 3. Placebo Test Distributions: Coefficient and t-statistics. This figure presents the distributions of the estimated coefficients (top panel) and corresponding *t*-statistics (bottom panel) from 1,000 placebo regressions. In each placebo iteration, we randomly reassign a fake launch date to each repository, sampling without replacement from the original set of launch dates to preserve the empirical distribution of event timing. Using these placebo event dates, we recalculate abnormal TVL(log) within the [−14, +14] event window and estimate the same regression as in Table3. We collect the estimated coefficient and *t*-statistic on the post-event indicator from each placebo regression. The red dashed vertical line denotes the actual coefficient (or *t*-statistic) from the main specification using true launch dates.

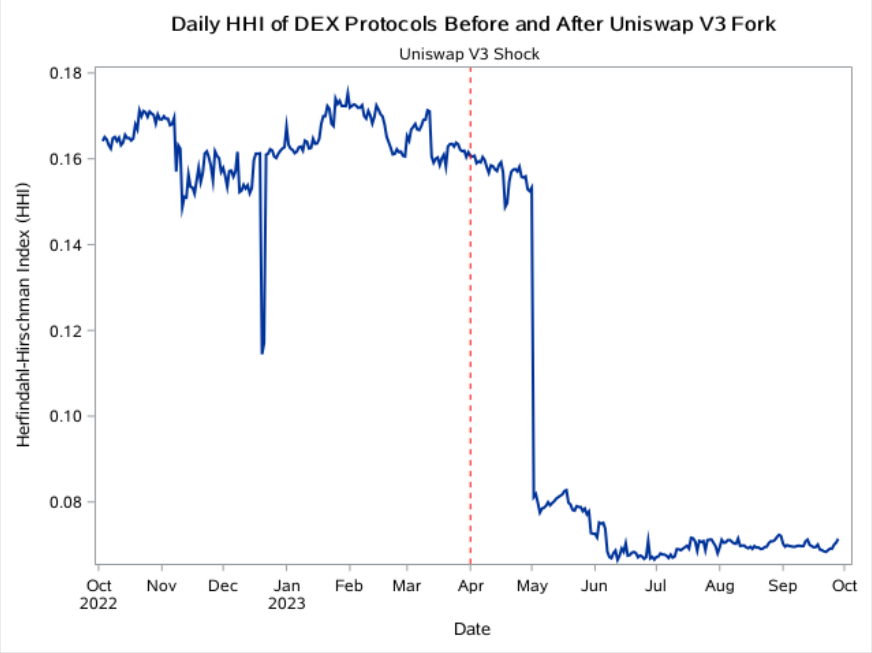


(a) Placebo Distribution of Estimated Coefficients

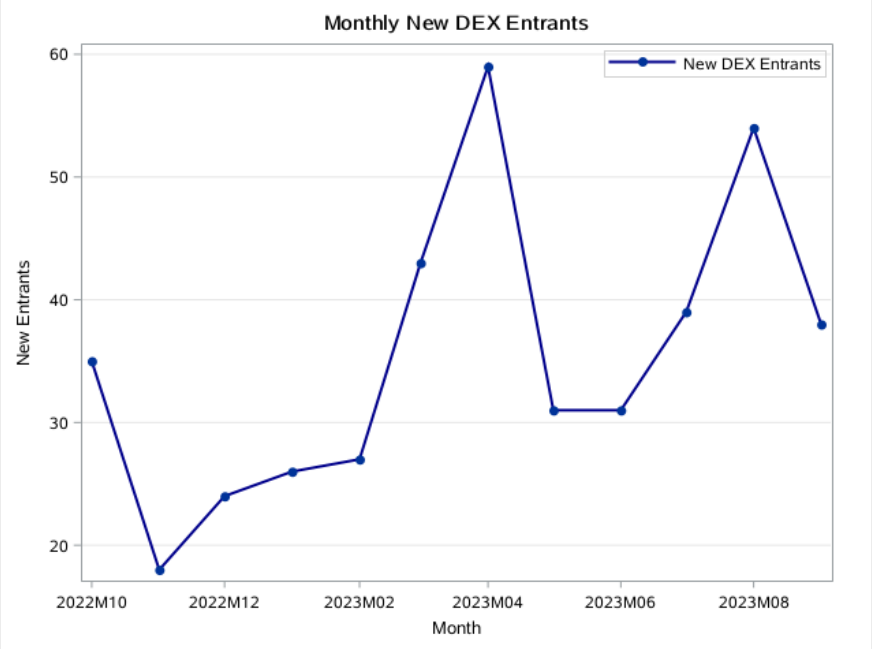


(b) Placebo Distribution of t-Statistics

Figure 4. Market Concentration of DEX Protocols Around the Uniswap V3 Copycat Wave. Panel (a) plots the evolution of market concentration among decentralized exchange (DEX) protocols, measured by the Herfindahl-Hirschman Index (HHI). Each protocol's daily market share is calculated as the ratio of its Total Value Locked (TVL) to the total TVL of all DEX protocols. The HHI is computed as the sum of squared market shares across all active protocols on each day, with higher values indicating greater concentration. Panel (b) shows the number of new DEX entrants launched each month. The red dashed vertical line marks the expiration of Uniswap V3's Business Source License (BSL) in April 2023, which triggered a wave of copycat protocols.



(a) Monthly HHI of the DEX Industry



(b) Monthly New Entrants in the DEX Industry

Figure 5. Parallel Trend: DEX Protocols vs. Non-DEX Protocols around the Uniswap V3 Copy Event. This figure illustrates the dynamic treatment effects of being a decentralized exchange (DEX) protocol relative to non-DEX protocols on the performance of newly launched repositories by DeFi protocols, measured over the monthly event window [−6, +6] around the Uniswap V3 license expiration in April 2023. The dependent variable is the logarithm of cumulative abnormal TVL over a 14-day window. All specifications include protocol characteristics and include industry, blockchain, and month fixed effects. Standard errors are clustered at the protocol industry level and reported in parentheses. Confidence intervals are shown at the 95% level.

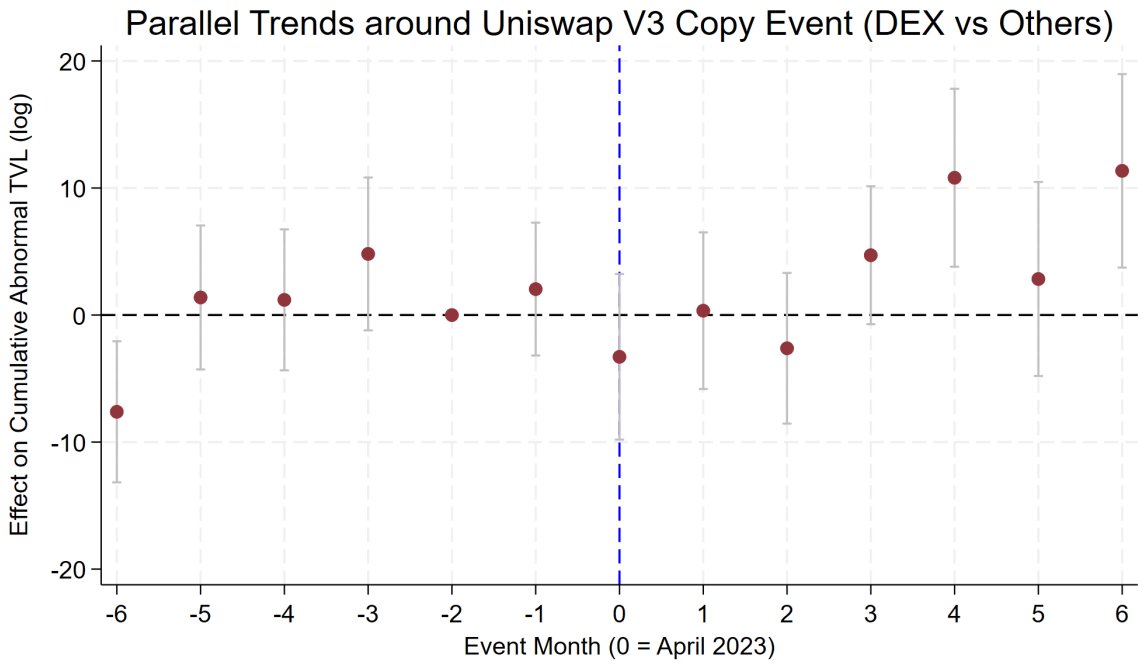


Figure 6. Quarterly DeFi Hack Losses and Hack Event Counts. This figure visualizes quarterly trends in decentralized finance (DeFi) security incidents from 2020 to 2024. The bars represent the total U.S. dollar value (in millions) lost due to smart contract exploits, cross-chain bridge breaches, and other protocol-level attacks. The grey line denotes the number of individual hack events occurring in each quarter. Details of the events are provided in Appendix C.

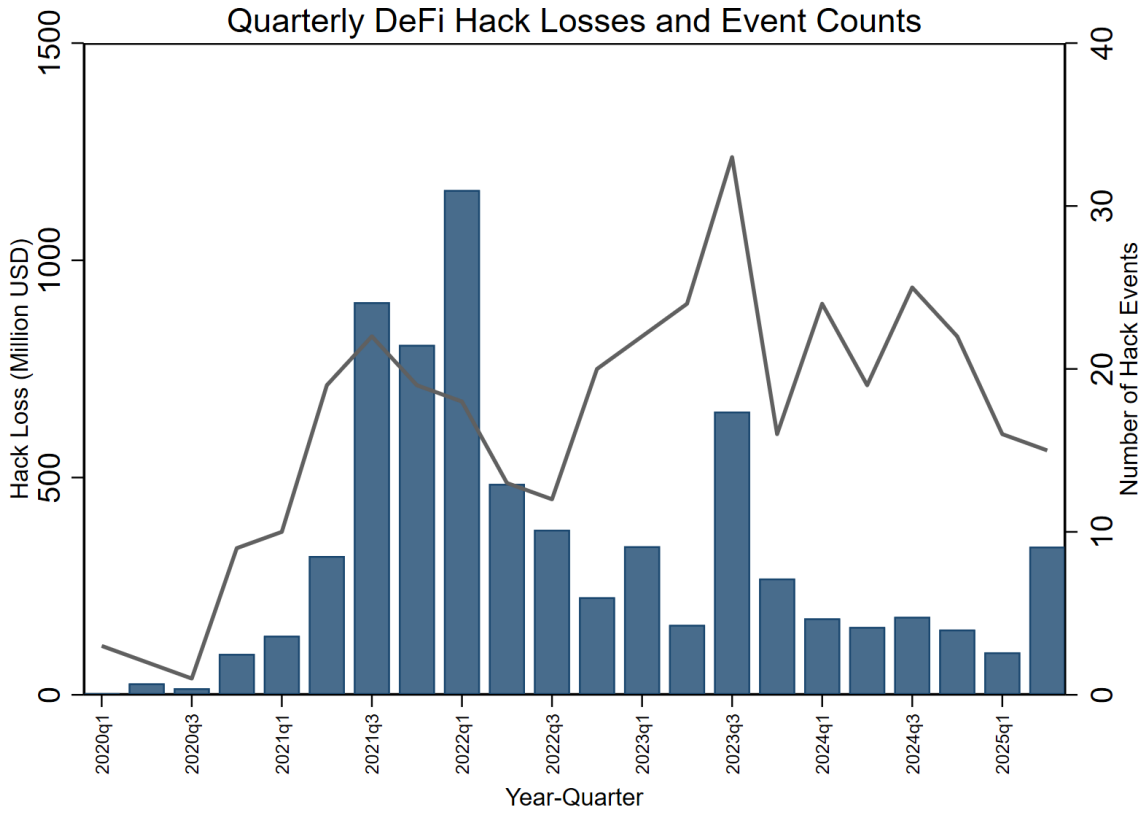


Table 1. Summary statistics. This table reports the summary statistics. Panel A presents statistics for the DeFi protocol-level sample, including the number of observations, mean, standard deviation, and interquartile range (25th, 50th, and 75th percentiles). Panel B reports the corresponding statistics for the repository-level sample. The main variables are described as below. *Size* is defined as the natural logarithm of a protocol's Total Value Locked (TVL) as of the GitHub repository's creation date. *AUDIT* is an indicator variable equal to one if a DeFi protocol's smart contracts have been audited by at least one auditor. *ORACLE* is an indicator variable equal to one if a protocol integrates external data feeds (e.g., asset prices) via oracle infrastructure. *LISTED* is an indicator variable that equals to one if the protocol has issued at least one tradable token on an exchange. *DAO* is an indicator variable equal to one if the protocol is governed by a Decentralized Autonomous Organization (DAO). *Lending* is an indicator variable that equals to one if a protocol whose primary functionality centers around lending and borrowing services. *Venture capital* is an indicator variable equal to one if the protocol used to receive early-stage venture capital(VC) funding. *Twitter followers (log)* is defined as the log-transformed number of followers of the protocol's official Twitter account. *Chains (log)* is defined as the natural logarithm of blockchains on which a protocol's smart contracts are deployed. *Staking (log)* is defined as the natural logarithm of assets staked within the protocol, capturing user engagement in governance or network security. *Popularity* is defined as the repository's stargazer count on GitHub. *Adoption* is measured by the number of forks a repository received, capturing downstream usage or adaptation of the code. *Maintenance* is defined as the number of unsolved issues released on GitHub. *Developer commit productivity* is defined as the average number of code commits by a repository's developers. *Developer team size* is defined as the number of distinct developers associated with the protocol before the repository launch. *Developer network* is defined as the average number of DeFi platforms each developer is affiliated with at the time of launch, reflecting cross-platform expertise.

Panel A. DeFi ventures						
	N	**mean**	**sd**	**p25**	**p50**	**p75**
SIZE	3,900	15.620	4.008	12.980	16.560	18.090
AUDIT	3,900	0.737	0.440	0	1	1
ORACLE	3,900	0.271	0.444	0	0	1
LISTED	3,900	0.470	0.499	0	0	1
Lending	3,900	0.118	0.322	0	0	0
DAO	3,900	0.163	0.369	0	0	0
Venture capital	3,900	0.246	0.431	0	0	0
Twitter followers (log)	3,900	7.002	4.141	5.493	8.339	10.040
Chains (log)	3,900	1.126	0.610	0.693	0.693	1.386
Staking (log)	3,900	4.227	6.593	0	0	10.490
Panel B. Open-source repositories						
Popularity (stargazer)	3,900	4.694	29.710	0	0	2
Extensibility(forks)	3,900	4.800	23.530	0	0	2
Analytics-oriented	3,900	0.038	0.192	0	0	1
Developer team size	3,900	1.911	5.210	0	0	1
Developer commit productivity	3,900	792.500	5,330	0	0	104
Developer network (cross-platform engagement)	3,900	5.455	16.520	0	0	1.571

Table 2. Event Study: Open-Source Repository Launches and Cumulative Abnormal Total Value Locked. This table presents the cumulative abnormal Total Value Locked (TVL) surrounding the launch of open-source repositories by DeFi protocols. Panel A reports mean-adjusted abnormal TVL, following Amiram et al. (2022). We first compute daily abnormal TVL as the difference between actual TVL and the mean TVL over a [-35, -16] pre-event window, and then aggregate it within each symmetric event window. Panel B presents market-adjusted abnormal TVL. For each protocol, we estimate expected TVL by regressing daily log TVL on the Ethereum market log TVL over the [-90, -15] window, and compute daily abnormal TVL as the residual. The outcome variable in both panels is the log cumulative abnormal TVL. We separately report estimates for all repository launch events and for first-time repository launches, defined as the first public GitHub repository disclosed by each protocol. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Mean-adjusted *cumulative abnormal TVL(log)*									
window	**obs.**	**[0]**	**[-1,+1]**	**[-2,+2]**	**[-3,+3]**	**[-4,+4]**	**[-5,+5]**	**[-6,+6]**	**[-7,+7]**
All repository launch events:									
Mean	4,299	0.044***	0.136***	0.224***	0.322***	0.403***	0.493***	0.587***	0.675***
t-values		(4.28)	(4.47)	(4.41)	(4.54)	(4.44)	(4.47)	(4.47)	(4.50)
First repository launch events:									
Mean	179	0.243***	0.732***	1.215***	1.873***	2.419***	2.979***	3.521***	4.064***
t-values		(2.68)	(2.67)	(2.65)	(2.99)	(3.00)	(3.02)	(3.01)	(3.02)
Panel B. Market-adjusted *cumulative abnormal TVL(log)*									
window	**obs.**	**[0]**	**[-1,+1]**	**[-2,+2]**	**[-3,+3]**	**[-4,+4]**	**[-5,+5]**	**[-6,+6]**	**[-7,+7]**
All repository launch events:									
Mean	4,298	0.021*	0.065**	0.107**	0.140*	0.170*	0.207*	0.252*	0.294*
t-values		(1.86)	(2.01)	(1.97)	(1.87)	(1.78)	(1.78)	(1.85)	(1.88)
First repository launch events:									
Mean	176	0.179*	0.555**	0.866*	1.265**	1.656**	2.062**	2.498**	2.947**
t-values		(1.96)	(2.02)	(1.93)	(2.03)	(2.09)	(2.16)	(2.24)	(2.31)

Table 3. Panel regression: The impact of Open-Source Repository Launches on Abnormal Total Value Locked (TVL). This table presents OLS regression results estimating the effect of new open-source repository launches on the abnormal Total Value Locked (TVL) of DeFi platforms. The dependent variable *Abnormal TVL (log)* is defined as the difference between the logarithm of a DeFi platform's actual TVL and the average logarithm TVL over the 20 trading days spanning from day -35 to day -16 prior to the event date. The sample includes daily observations within symmetric windows of [-3,+3], [-5,+5], [-7,+7], and [-14,+14] surrounding each event. *POST* is a dummy variable equal to one for post-event observations ($t \geq 0$), where $t = 0$ denotes the launch date of the new open-source repository. Repository-level fixed effects are included in Columns (1), (3), (5), and (7), whereas Columns (2), (4), (6), and (8) include industry and blockchain fixed effects instead. All specifications include fixed effects for repository launch month. Standard errors are clustered at the repository level. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Abnormal TVL (log)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WINDOW	**[-3,+3]**	**[-3,+3]**	**[-5,+5]**	**[-5,+5]**	**[-7,+7]**	**[-7,+7]**	**[-14,+14]**	**[-14,+14]**
POST (New repository)	0.008**	0.008**	0.008**	0.009**	0.009**	0.010**	0.014**	0.016**
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)	(0.006)
SIZE		0.006		0.006		0.006		0.006*
		(0.004)		(0.004)		(0.004)		(0.004)
Chains (log)		0.071*		0.063		0.061		0.064*
		(0.040)		(0.040)		(0.039)		(0.038)
Twitter followers (log)		0.001		0.001		0.001		0.001
		(0.004)		(0.004)		(0.004)		(0.003)
ORACLE		-0.024		-0.028		-0.027		-0.020
		(0.045)		(0.044)		(0.044)		(0.042)
DAO		0.052		0.054		0.056*		0.058*
		(0.035)		(0.035)		(0.034)		(0.031)
Venture capital		0.110***		0.115***		0.117***		0.119***
		(0.036)		(0.035)		(0.035)		(0.033)
Staking (log)		-0.003		-0.003		-0.003		-0.003*
		(0.002)		(0.002)		(0.002)		(0.002)
AUDIT		-0.058*		-0.053*		-0.053*		-0.055**
		(0.030)		(0.030)		(0.029)		(0.028)
LISTED		-0.115***		-0.115***		-0.117***		-0.114***
		(0.026)		(0.025)		(0.025)		(0.023)
Constant		-0.053		-0.058		-0.054		-0.074
		(0.081)		(0.079)		(0.078)		(0.072)
Observations	30,205	27,419	47,404	43,029	64,598	58,641	124,650	113,168
R-squared	0.952	0.149	0.929	0.147	0.909	0.146	0.819	0.139
Repo FE	YES	NO	YES	NO	YES	NO	YES	NO
Industry FE	NO	YES	NO	YES	NO	YES	NO	YES
Blockchain FE	NO	YES	NO	YES	NO	YES	NO	YES
Year-mon FE	YES	YES	YES	YES	YES	YES	YES	YES

Table 4. DiD analysis: The Effect of Transitioning from Closed-Source to Open-Source on Abnormal TVL. This table reports results from difference-in-differences regressions estimating the effect of transitioning from closed-source to open-source on DeFi protocols' abnormal Total Value Locked (TVL), measured in logarithmic form. The treatment group consists of protocols that launched their first public GitHub repository—representing a transition to open-source—while the control group includes protocols that remained closed-source throughout the sample period. Control protocols are assigned a pseudo launch date based on their matched treated protocol's actual repository launch date. We implement 1:1 nearest-neighbor propensity score matching based on protocol characteristics including *SIZE*, *Chains(log)*, *Twitter followers(log)*, *ORACLE*, *DAO*, *Venture capital*, *Staking(log)*, *AUDIT*, and *LISTED*. Each control protocol is assigned a pseudo event date corresponding to the matched treated protocol's actual open-source launch. The sample includes 14 days before and after the (actual or pseudo) repository launch date. The variable *POST* is an indicator for days on or after the launch date. All regressions control for protocol-level characteristics and include fixed effects for industry, blockchain, and year-month. Standard errors are clustered at the industry level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ***Abnormal TVL (log)***	(2) ***Abnormal TVL (log)***
WINDOW	**[-14, +14]**	**[-14, +14]**
Treated (OpenSource)		0.271**
		(0.108)
*Treated (OpenSource) * POST (New repository)*	0.162*	0.160*
	(0.085)	(0.083)
Observations	6,329	6,329
R-squared	0.866	0.396
Protocol FE	YES	NO
Industry FE	NO	YES
Blockchain FE	NO	YES
Year-mon FE	YES	YES

Table 5. Two-stage least squares (2SLS) regression with the instrument of GitHub Copilot AI Tooling Releases. This table reports the results of an instrumental variable (IV) strategy to estimate the causal impact of open-source repository launches on the Total Value Locked (TVL) of DeFi protocols. The instrument is based on staggered releases of GitHub Copilot-related AI tooling, including Copilot Preview (2021-06), Copilot General Release (2022-06), Codespaces Free Tier (2022-11), Copilot X Preview (2023-03), and Copilot Chat Release (2023-12), as illustrated in Figure 1. We define the instrument variable, *GitHub Copilot Launch*, as an indicator equal to 1 if a protocol-month falls within one month following any of the GitHub Copilot-related AI tooling releases. The first stage models the probability that a new GitHub repository is launched by a DeFi protocol in a given month. The second stage regresses the log of Total Value Locked (TVL) on the predicted probability of new repository creation. TVL is defined as the total value locked in the protocol for each month. All regressions control for protocol-level characteristics and include fixed effects for industry and blockchain. Standard errors clustered at the protocol level are reported in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	First-stage	**Second-stage**	
	(1)	(2)	(3)
	New Repository	TVL_t ***(log)***	TVL_{t+1} ***(log)***
IV = *GitHub Copilot Launch*	0.004***		
	(0.001)		
New Repository		62.454***	236.390***
		(22.265)	(71.827)
Observations	90,798	90,783	90,798
R-squared	0.044	0.039	0.051
Control	YES	YES	YES
Industry FE	YES	YES	YES
Blockchain FE	YES	YES	YES
First-stage partial F-test	10.40		

Table 6. Open Source and Mimicking Risk: Evidence from Uniswap V3 License Expiration. This table evaluates the effect of the Uniswap V3 license expiration (April 2023) on the performance of newly launched repositories by DeFi protocols. The treated group includes protocols in the decentralized exchange (DEX) sector, while the control group consists of protocols from other sectors (e.g., Lending, Yield, Derivatives, etc.). The sample includes all protocols operating on multiple blockchains that launched a new GitHub repository within six months before or after the Uniswap V3 copycat wave. Panel A reports the impact on cumulative abnormal Total Value Locked (TVL), measured in natural logarithm, over three symmetric windows around each repository launch: [-1,+1], [-7,+7], and [-14,+14] days. Panel B presents results for market-adjusted cumulative abnormal token returns over the same windows. The key independent variable is the interaction term *Treated (DEX)* × *Post Uniswap*, where the post period begins in April 2023, following the expiration of Uniswap V3's Business Source License. All regressions control for protocol characteristics and include industry, blockchain, and year-month fixed effects. Standard errors are clustered at the protocol industry level and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Cumulative abnormal TVL (log)			
	(1)	(2)	(3)
WINDOW	**[-1, +1]**	**[-7,+7]**	**[-14,+14]**
Treated (DEX) × *Post Uniswap*	0.862***	4.873***	10.835***
	(0.245)	(1.197)	(2.369)
Observations	561	561	561
R-squared	0.165	0.148	0.158
Industry FE	YES	YES	YES
Blockchain FE	YES	YES	YES
Year-mon FE	YES	YES	YES
Panel B. Cumulative abnormal returns			
WINDOW	**[-1, +1]**	**[-7,+7]**	**[-14,+14]**
Treated (DEX) × *Post Uniswap*	0.040***	0.114**	0.124*
	(0.013)	(0.042)	(0.070)
Observations	461	461	461
R-squared	0.128	0.133	0.192
Industry FE	YES	YES	YES
Blockchain FE	YES	YES	YES
Year-mon FE	YES	YES	YES

Table 7. Open Source and Cybersecurity Risk: Evidence from DeFi Hack Events. This table reports the effect of open-source repository launches on the subsequent security outcomes of DeFi protocols, measured by hacking losses and attack probability. The treated group consists of protocols that launched their first public GitHub repository in a given year, while the control group includes protocols that never launched a repository throughout the sample period. We implement 1:1 nearest-neighbor propensity score matching based on protocol characteristics including *SIZE*, *Chains(log)*, *Twitter followers(log)*, *ORACLE*, *DAO*, *Venture capital*, *Staking(log)*, *AUDIT*, and *LISTED*. Column (1) presents a linear regression (OLS) where the dependent variable is a dummy equal to one if the protocol experienced at least one hack event in the same post-event window. Column (2) reports a linear regression (OLS) where the dependent variable is *HACK loss(log)*, measured as the log of total hacking losses (in USD) experienced by the protocol following the event year. All specifications include protocol-level controls, industry fixed effects, blockchain fixed effects, and year fixed effects. Standard errors are clustered at the protocol industry level and reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ***HACK Dummy***	(2) ***HACK Loss (log)***
Treated	0.046**	0.768*
	(0.021)	(0.370)
Observations	562	562
R-squared	0.181	0.188
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year FE	YES	YES

Table 8. Event study for new open-source repository launches: cumulative abnormal returns (Token Price). This table presents the cumulative abnormal return (CAR) responses of tokens to the launch of new open-source repositories, based on a market-adjusted return model. Daily abnormal returns are calculated as the difference between the actual token return and the expected return derived from Ethereum market returns, estimated using a regression over the [-90, -15] pre-event window, following Liu and Tsyvinski (2021). Panel A reports short-term CARs within event windows of [0], [-1,+1], [-2,+2], and [-3,+3]. Panel B reports longer-term CARs over [0,90], [0,120], [0,150], and [0,180] day windows. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Short-term: Market-adjusted Culmulative abnormal returns					
window	**Obs.**	**[0]**	**[-1,+1]**	**[-2,+2]**	**[-3,+3]**
Mean	3,640	-0.0080***	-0.0096***	-0.0159***	-0.0057
t values		(-3.70)	(-3.13)	(-3.66)	(-1.30)
Panel B. Long-term: Market-adjusted Culmulative abnormal returns					
window		**[0, 90]**	**[0, 120]**	**[0, 150]**	**[0, 180]**
Mean	3,640	0.079*	0.133**	0.147**	0.179**
t values		(1.84)	(2.17)	(2.16)	(2.33)

Table 9. Long-Term Product Innovation Outcomes: Development of New Features in DeFi Protocols. This table examines the long-run impact of open-source code disclosure on product innovation in DeFi protocols. The dependent variable is *New Features (log)*, defined as the logarithm of the number of newly introduced functional modules or capabilities—such as staking, lending, token bridging, or governance upgrades—by each protocol in a given year. The treated group consists of protocols that launched their first public GitHub repository during the sample period, while the control group includes protocols that never launched a repository. We apply 1:1 nearest-neighbor propensity score matching based on protocol characteristics, including *Size*, *Chains (log)*, *Twitter Followers (log)*, *Oracle*, *DAO*, *Venture Capital*, *Staking (log)*, *Audit*, and *Listed*. Each matched control protocol is assigned the same event year as its corresponding treated protocol. The variable *Post* is an indicator equal to one if the year is equal to or after the protocol's repository launch year, and zero otherwise. Column (1) reports results from a standard difference-in-differences (DID) specification. Column (2) presents a dynamic DID analysis, using the repository launch year as the omitted category. All regressions include industry, blockchain, and calendar year fixed effects. Standard errors clustered at the protocol level are shown in parentheses. *, **, and * * * indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1) ***New features (log)***	(2) ***New features (log)***
Treated × *Post*	0.328***	
	(0.120)	
Treated × $Before^{2yr}$		0.106
		(0.085)
Treated × $Before^{1yr}$		0.220
		(0.136)
Treated × $After^{1yr}$		1.149***
		(0.318)
Treated × $After^{2yr}$		0.521**
		(0.211)
Treated	−0.204**	−0.291***
	(0.088)	(0.077)
Observations	2,350	2,350
R-squared	0.317	0.328
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year FE	YES	YES

Table 10. Governance and Monitoring Channel: Protocol-Level Organizational Characteristics. This table investigates how the effect of open-source code disclosure on abnormal Total Value Locked (TVL) varies with protocols' organizational features. We examine three dimensions: (i) governance structure, proxied by the presence of decentralized autonomous organizations (DAOs, Panel A), where *DAOs* adopt token-based governance structures, and *Non-DAOs* do not; (ii) external monitoring, captured by venture capital backing (Panel B), where *VC-backed* protocols have identifiable VC investors, while *Non-VC-backed* do not; and (iii) user-driven oversight incentives, proxied by participation in the lending sector (Panel C), which involves greater user asset exposure and monitoring demand. The dependent variable is the log of abnormal TVL over the [−14, +14] day window around repository disclosure. All regressions include protocol-level controls and fixed effects for industry, blockchain, and year-month. *, **, and * * * indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Governance structure		
	Abnormal TVL (log)	***Abnormal TVL (log)***
	(1)	(2)
WINDOW	**[-14,+14]**	**[-14,+14]**
	DAOs	**Non-DAOs**
POST (New repository)	0.053***	0.008
	(0.018)	(0.007)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	18,412	94,756
R-squared	0.193	0.181
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel B. External monitor		
	VC-backed	**Non-VC-backed**
POST (New repository)	0.051***	0.005
	(0.009)	(0.008)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	27,890	85,278
R-squared	0.297	0.146
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel C. User risk exposure		
	Lending protocols	**Other industries**
POST (New repository)	0.069**	0.009
	(0.027)	(0.006)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	13,326	99,842
R-squared	0.368	0.135
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES

Table 11. Innovation Signaling Channel: Repository-Level Innovation Quality. This table reports the effect of innovation quality of code disclosure on abnormal Total Value Locked (TVL). We categorize repositories into three dimensions of innovation quality: (i) popularity, measured by the number of GitHub stargazers (Panel A), where *HIGH Stargazers* refers to repositories in the top quartile of the star distribution, and *LOW Stargazers* includes the remaining repositories; (ii) extensibility, measured by the number of forks (Panel B), where *HIGH Forks* refers to repositories in the top quartile of the fork distribution; and (iii) transparency, proxied by whether the repository is analytics-oriented (Panel C). A repository is defined as analytics-oriented based on textual analysis of metadata fields, identifying references to data reporting, dashboard functionality, or user-facing metrics. The dependent variable is the log of abnormal TVL over a [−14, +14] day window surrounding the repository disclosure date. All regressions include protocol-level controls as well as industry, blockchain, and year-month fixed effects. ∗, ∗∗, and ∗ ∗ ∗ denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Popularity of repository		
	Abnormal TVL (log)	***Abnormal TVL (log)***
	(1)	(2)
WINDOW	**[-14,+14]**	**[-14,+14]**
	HIGH Stargazers	**LOW Stargazers**
POST (New repository)	0.025**	0.009
	(0.010)	(0.009)
Diff.(test)	Prob > $\chi^2(1)$ = 0.045	
Observations	45,710	67,458
R-squared	0.191	0.150
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel B. Extensibility of repository		
	HIGH Forks	**LOW Forks**
POST (New repository)	0.023***	0.009
	(0.008)	(0.010)
Diff.(test)	Prob > $\chi^2(1)$ = 0.049	
Observations	51,257	61,911
R-squared	0.228	0.100
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel C. Transparency of repository		
	Analytics-oriented	**Non-analytics-oriented**
POST (New repository)	0.057*	0.014**
	(0.034)	(0.007)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	4,094	109,074
R-squared	0.617	0.138
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES

Table 12. Development Engagement and Network Channel: Developer-team level Characteristics This table explores how the impact of open-source code disclosure on abnormal Total Value Locked (TVL) varies with developer team characteristics. We classify repositories based on three dimensions: (i) team size, measured by the total number of distinct developers contributing to the repository (Panel A); (ii) developer productivity, measured as the average number of commits per developer (Panel B); and (iii) developer network connectivity, defined as the average number of distinct repositories each developer has contributed to across the platform (Panel C). All subgroup splits are based on the top quartile (High) versus the remaining distribution (Low). The dependent variable is the log of abnormal TVL over the [−14, +14] day window surrounding the disclosure. All regressions include protocol-level controls and fixed effects for industry, blockchain, and year-month. *, **, and * * * indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Developer team size		
	Abnormal TVL (log)	***Abnormal TVL (log)***
	(1)	(2)
WINDOW	**[-14,+14]**	**[-14,+14]**
	HIGH Size	**LOW Size**
POST (New repository)	0.039**	0.008
	(0.017)	(0.007)
Diff.(test)	Prob > $\chi^2(1)$ = 0.002	
Observations	26,308	86,860
R-squared	0.135	0.191
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel B. Productivity of developers		
	HIGH Productivity	**LOW Productivity**
POST (New repository)	0.073***	—0.002
	(0.018)	(0.006)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	26,255	86,913
R-squared	0.213	0.140
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES
Panel C. Network of developers		
	HIGH Network	**LOW Network**
POST (New repository)	0.076***	0.000
	(0.020)	(0.006)
Diff.(test)	Prob > $\chi^2(1)$ = 0.000	
Observations	22,618	90,550
R-squared	0.157	0.183
Control	YES	YES
Industry FE	YES	YES
Blockchain FE	YES	YES
Year-mon FE	YES	YES

Appendix A: Figures and Tables

Figure A1. Evolution of Total Value Locked (TVL). This figure plots the time-series evolution of the logarithm of Total Value Locked (TVL) across DeFi protocols from 2019 to 2024.

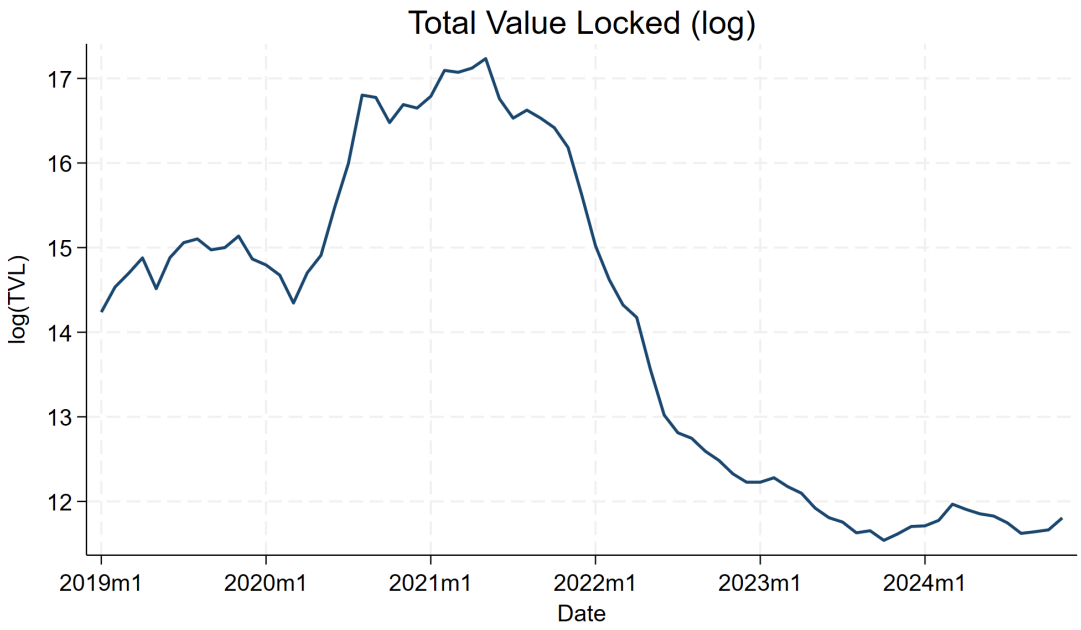
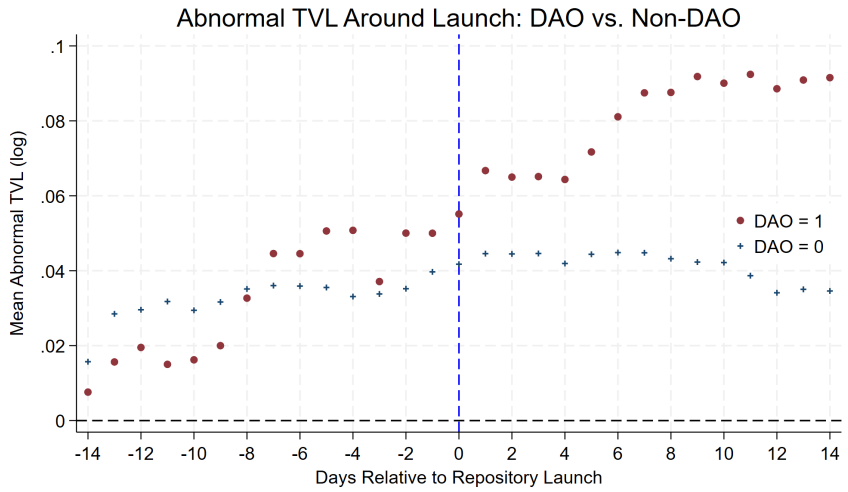
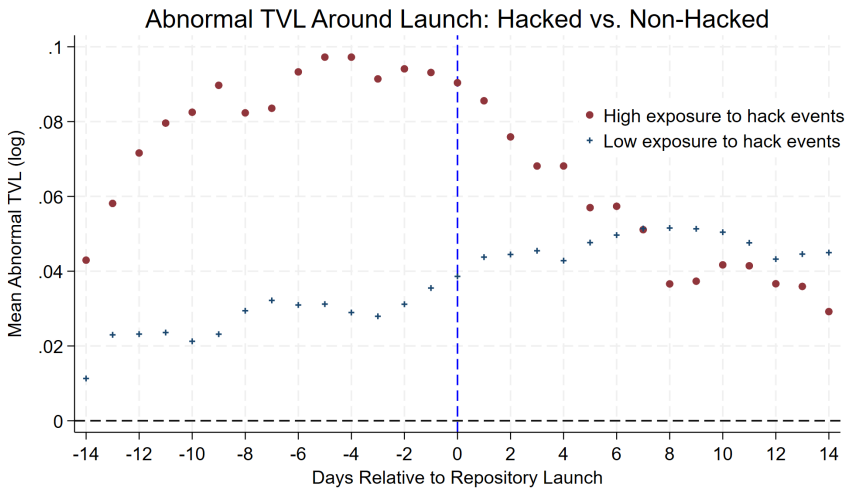


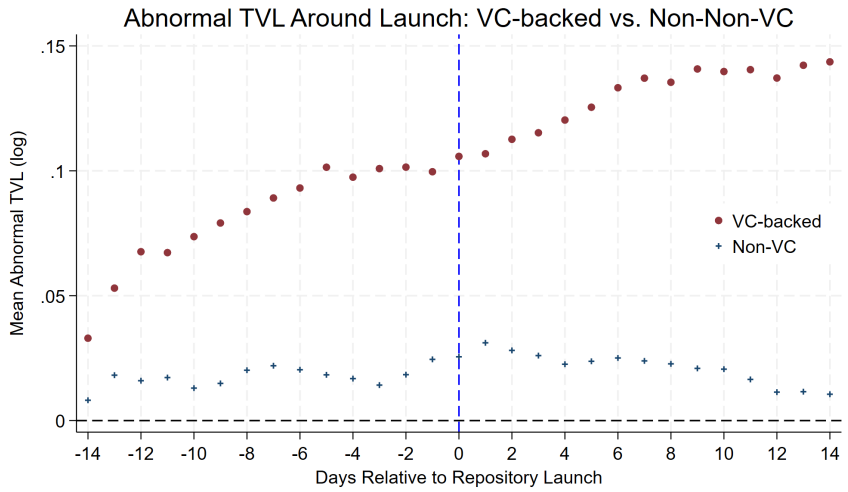
Figure A2. Abnormal TVL around new repositories launched: heterogeneous responses by protocol characteristics. This figure plots the daily abnormal Total Value Locked (TVL) surrounding new open-source repository launches over the [-14, +14] event window relative to the launch date. Panels (a) through (d) present subsample comparisons based on governance structure, security risk, investor backing, and protocol type, respectively. Panel (a) compares protocols with decentralized autonomous organization (DAO) governance to those without. Panel (b) compares protocols with high exposure to hack events to those without. Panel (c) compares protocols backed by venture capital (VC) to those without VC support. Panel (d) compares lending protocols to protocols in other sectors.



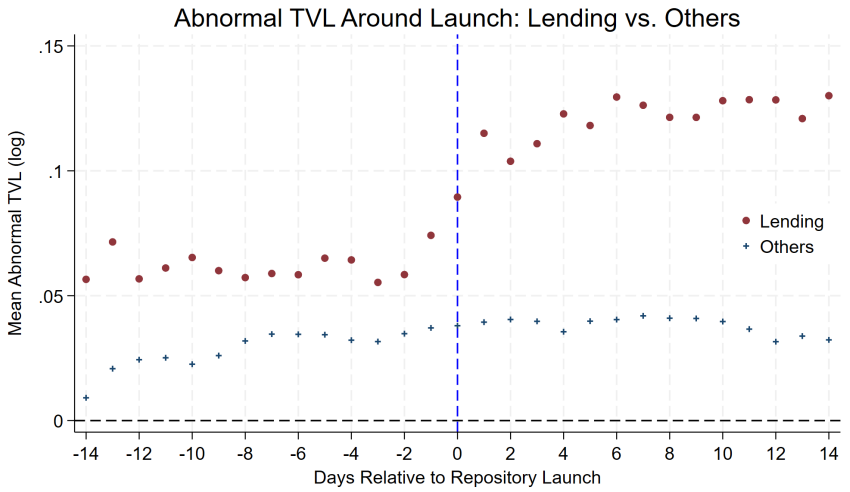
(a) DAO vs. Non-DAO



(b) High exposure to hack events vs. Low exposure to hack events



(c) VC-backed vs. Non-VC-backed



(d) Lending protocols vs. Other Industries

Table A1. Definitions of the variables used in the analysis.

Variable	Definition
TVL	the total dollar value of digital assets locked in a DeFi protocol (in USD).
Token Returns	the log difference in token price over a daily interval, calculated as $\log(p_t) - \log(p_{t-1})$, where p_t is the token's closing price in USD on day t .
AUDIT	an indicator variable equal to one if a DeFi protocol's smart contracts have been audited by at least one auditor.
ORACLE	an indicator variable equal to one if a protocol integrates external data feeds (e.g., asset prices) via oracle infrastructure.
SIZE	the natural logarithm of a protocol's Total Value Locked (TVL) as of the GitHub repository's creation date.
LISTED	an indicator variable that equals to one if the protocol has issued at least one tradable token on an exchange.
Lending	an indicator variable that equals to one if a protocol whose primary functionality centers around lending and borrowing services.
Twitter followers (log)	the log-transformed number of followers of the protocol's official Twitter account.
Chains (log)	the natural logarithm of blockchains on which a protocol's smart contracts are deployed.
Staking (log)	the natural logarithm of assets staked within the protocol, capturing user engagement in governance or network security.
DAO	an indicator variable equal to one if the protocol is governed by a Decentralized Autonomous Organization (DAO).
Venture capital	an indicator variable equal to one if the protocol used to receive early-stage venture capital(VC) funding.
HACK	an indicator variable that equals to one if the repository launch month coincides with one or more documented DeFi security breaches.
SIZE	the natural logarithm of a protocol's Total Value Locked (TVL) as of the GitHub repository's creation date.
Popularity (stargazers)	the repository's stargazer count on GitHub.
Extensibility (forks)	the number of forks a repository received, capturing adaptation of the code.
Transparency(analytics-oriented)	an indicator variable equal to one if the repository metadata includes terms related to data analytics, such as reporting tools, dashboards, or user metrics, as identified through keyword-based text analysis.
Developer team size	the number of distinct developers associated with the protocol before the repository launch.
Developer productivity	the average number of code commits by a repository's developers.
Developer network	the average number of DeFi platforms each developer is affiliated with at the time of launch, reflecting cross-platform expertise.

Table A2. Industry categories: Business type and cumulative abnormal TVL. This table reports cumulative abnormal Total Value Locked (TVL) around new open-source repository launches, categorized by business type. We classify DeFi startups into three main groups: Lending platforms, decentralized exchanges (DEX), and others. We compute mean-adjusted and market-adjusted cumulative abnormal TVL for each group over the event window. All TVL measures are computed in logarithmic form.

Industry	Obs.	Mean-adjusted cumulative abnormal TVL		Market-adjusted cumulative abnormal TVL	
		[0]	[-1,+1]	[0]	[-1,+1]
Lending	483	0.089**	0.299**	0.124**	0.383**
		(2.30)	(2.64)	(2.97)	(3.13)
DEX	1,477	0.021	0.062	—0.010	—0.025
		(1.15)	(1.15)	(—0.55)	(—0.45)
Others	2,333	0.048**	0.149**	0.019	0.057
		(3.89)	(3.98)	(1.31)	(1.37)

Table A3. Summary Statistics: Propensitory score matching sample. This table reports the summary statistics of protocol characteristics for open-source and closed-source startups, before and after 1:1 propensity score matching. The treated group comprises protocols that launched a public GitHub repository, while the control group includes those that did not. Matching is performed using 1:1 nearest-neighbor matching without replacement based on a logit model that includes the following covariates: *size*, *log(Chains)*, *ORACLE*, *DAO*, *Venture Capital*, *log(Staking)*, *LISTED*, *AUDIT*, and industry fixed effects (e.g., Lending). Columns (1)-(3) display the means and standard deviations for unmatched samples, while columns (4)-(6) show those for the matched sample. For each variable, we report the average value in the treated (open-source) and control (closed-source) groups, as well as the mean difference between groups and the associated t-test statistic. Standard deviations are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. All TVL-related variables are measured in logarithmic form.

	Before-matching sample			**After 1:1 matching sample**		
	(1) **Close-source**	(2) **Open-source**	(3) **Difference**	(4) **Close-source**	(5) **Open-source**	(6) **Difference**
Size	11.137	13.759	2.622***	13.894	13.759	-0.135
	(5.377)	(3.765)	(0.323)	(3.498)	(3.765)	(0.303)
Chains(log)	0.874	1.050	0.176***	1.030	1.074	0.044
	(0.387)	(0.522)	(0.022)	(0.497)	(0.543)	(0.043)
ORACLE	0.171	0.262	0.091***	0.279	0.289	0.010
	(0.376)	(0.440)	(0.021)	(0.449)	(0.454)	(0.038)
DAO	0.084	0.215	0.130***	0.226	0.216	-0.010
	(0.278)	(0.411)	(0.016)	(0.419)	(0.412)	(0.035)
Venture Capital	0.046	0.150	0.104***	0.139	0.146	0.007
	(0.211)	(0.358)	(0.013)	(0.347)	(0.354)	(0.029)
Staking(log)	1.762	5.278	3.516***	4.900	4.638	-0.262
	(4.227)	(6.520)	(0.250)	(6.376)	(6.230)	(0.526)
LISTED	0.108	0.535	0.427***	0.463	0.481	0.017
	(0.311)	(0.499)	(0.018)	(0.500)	(0.501)	(0.042)
AUDIT	0.441	0.685	0.244***	0.749	0.728	-0.021
	(0.497)	(0.465)	(0.028)	(0.434)	(0.446)	(0.037)
Lending	0.095	0.144	0.049***	0.150	0.157	0.007
	(0.293)	(0.352)	(0.017)	(0.358)	(0.364)	(0.030)
Observations	4,324	340	4,664	287	287	574

Table A4. Instrument Validity: Exclusion Restriction Test. This table reports the results from regressions examining whether the instrumental variable, *GitHub Copilot Launch*, is directly associated with platform outcomes. The dependent variable is the log of Total Value Locked (TVL), measured at the protocol-day level. *GitHub Copilot Launch* captures the timing of Copilot-related tool launches, which is used in the main analysis as an instrument for open-source repository disclosure. Column (1) includes only the instrument variable and fixed effects, while column (2) adds a full set of control variables. Industry and blockchain fixed effects are included in all specifications. Standard errors are clustered at the protocol level and reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1) ***log(TVL)***	(2) ***log(TVL)***
IV = GitHub Copilot Launch	0.070	0.068
	(0.056)	(0.059)
Chains (log)		1.352***
		(0.197)
Twitter followers (log)		0.275***
		(0.031)
ORACLE		0.411**
		(0.160)
DAO		0.373*
		(0.197)
Venture capital		0.619**
		(0.260)
Staking (log)		0.092***
		(0.014)
AUDIT		1.174***
		(0.137)
LISTED		0.544***
		(0.199)
Observations	90,783	90,783
R-squared	0.190	0.303
Industry FE	YES	YES
Blockchain FE	YES	YES

Appendix B. Uniswap V3 Copycat Wave: Post-License Expiry Forks

Uniswap V3, launched in May 2021, introduced a significant innovation—concentrated liquidity—which enabled liquidity providers (LPs) to allocate capital more efficiently within specified price ranges. To safeguard its proprietary design, Uniswap Labs adopted a two-year Business Source License (BSL), restricting the use of its source code for commercial purposes. Following the BSL's expiration on April 1, 2023, the V3 code was re-released under the GNU General Public License (GPL), allowing unrestricted public use. This transition triggered a rapid wave of copycat deployments by competing protocols. PancakeSwap V3, for example, launched its fork just two days later on April 3, 2023 (pancakeswap github), attracting over $125 million in Total Value Locked (TVL) within days.[25] Other notable forks include QuickSwap V3 and Algebra, which deployed Uniswap V3-style pools across BNB Chain, Polygon, and Arbitrum.

In the immediate aftermath, Uniswap V3's own TVL declined modestly. According to DeFiLlama, its TVL fell from approximately $2.8 billion in late March 2023 to around $2.6 billion by mid-April—a reduction of roughly 7%. This contraction highlights the competitive pressure induced by open-source licensing expiration. In response, Uniswap Labs accelerated its own multichain deployments across Layer 1 and Layer 2 ecosystems to retain user liquidity and protocol dominance.

B1. Summary of Uniswap Copycat Events and Forks.

Event Type	Event Date	Copycat Protocols	Copy Mechanism
Uniswap V2 Copycat	2020-09-09	SushiSwap	Direct fork with added governance token (SUSHI) and LP incentives
Uniswap V3 Copycat	2023-04-01	PancakeSwap V3, Algebra, QuickSwap V3	Direct fork post-license expiry; multi-chain deployment

[25] See UnchainedCrypto, April 2023.

Appendix C. DeFi hack events

ID	Name	Date	ID	Name	Date	ID	Name	Date
1	Ronin Bridge	23 Mar, 2022	41	Bunny	19 May, 2021	81	Inverse Finance Frontier	2 Apr, 2022
2	Poly Network	10 Aug, 2021	42	Hedgey	19 Apr, 2024	82	Eminence	28 Sep, 2020
3	Binance Bridge	6 Oct, 2022	43	BingX	20 Sep, 2024	83	Rain	29 Apr, 2024
4	Portal	2 Feb, 2022	44	Stake.com	4 Sep, 2023	84	Furucombo	27 Feb, 2021
5	Ambient	17 Oct, 2024	45	Alpha Finance	13 Feb, 2021	85	DEUS Finance	28 Apr, 2022
6	Mixin Network	23 Sep, 2023	46	CoinsPaid	22 Jul, 2023	86	pNetwork	4 Oct, 2021
7	Euler V1	13 Mar, 2023	47	Vee Finance	21 Sep, 2021	87	Abracadabra Spell	25 Mar, 2025
8	Nomad	1 Aug, 2022	48	Meerkat Finance	4 Mar, 2021	88	Polter Finance	16 Nov, 2024
9	Beanstalk	17 Apr, 2022	49	MonoX	30 Nov, 2021	89	Ronin Bridge	6 Aug, 2024
10	Wintermute	20 Sep, 2022	50	Spartan	2 May, 2021	90	Cork V1	28 May, 2023
11	Parity Multisig	9 Nov, 2017	51	Grim Finance	18 Dec, 2021	91	Compounder Finance	2 Dec, 2020
12	Compound V2	29 Sep, 2021	52	Wintermute	5 Jun, 2022	92	PrismaLST	28 Mar, 2024
13	Vulcan Forged	13 Dec, 2021	53	StableMagnet	23 Jun, 2021	93	Yearn Finance	13 Apr, 2023
14	CREAM Lending	27 Oct, 2021	54	Paid Network	5 Mar, 2021	94	Saddle Finance	30 Apr, 2022
15	Multichain	7 Jul, 2023	55	Penpie	3 Sep, 2024	95	ValueDefi	7 May, 2021
16	Badger DAO	2 Dec, 2021	56	FixedFloat	18 Feb, 2024	96	Yearn Finance	5 Feb, 2021
17	Mango Markets V3	11 Oct, 2022	57	Kronos Research	18 Nov, 2023	97	Rari Capital	8 May, 2021
18	Atomic Wallet	3 Jun, 2023	58	ThalaSwap	15 Nov, 2024	98	Arbix Finance	4 Jan, 2022
19	Harmony Bridge	23 Jun, 2022	59	Indodax	10 Sep, 2024	99	Dego Finance	10 Feb, 2022
20	Mirror	8 Oct, 2021	60	Harvest Finance	26 Oct, 2020	100	ValueDefi	5 May, 2021
21	Heco Bridge	22 Nov, 2023	61	dForce	19 Apr, 2020	101	DogWifTools	28 Jan, 2025
22	Phemex	23 Jan, 2023	62	xToken	12 May, 2021	102	LiFi Finance	16 Jul, 2024
23	Harbor Protocol	19 Aug, 2023	63	Bald	2 Aug, 2023	103	zkLend	11 Feb, 2025
24	Orbit Bridge	31 Dec, 2023	64	Elephant Money	12 Apr, 2022	104	Cover Protocol	29 Dec, 2020
25	Qubit	28 Jan, 2022	65	Gala	20 May, 2024	105	MyAlgo	27 Feb, 2023
26	Rari Capital	1 May, 2022	66	Pando Rings	6 Nov, 2022	106	dYdX V3	18 Nov, 2023
27	AscendEX	11 Dec, 2021	67	Blizz Finance	13 May, 2022	107	Punk Protocol	10 Aug, 2021
28	Munchables	26 Mar, 2024	68	DEXX	16 Nov, 2024	108	Safemoon	28 Mar, 2023
29	Curve DEX	30 Jul, 2023	69	Transit Swap	2 Oct, 2022	109	Zoth ZeUSD	21 Mar, 2022
30	AlphaPo	26 Jul, 2023	70	Popsicle Finance	3 Aug, 2021	110	Crema Finance	2 Jul, 2022
31	The DAO	17 Jun, 2016	71	Sonne Finance	15 May, 2024	111	Superfluid	8 Feb, 2022
32	AnubisDAO	29 Oct, 2021	72	US Gov Crypto Wallet	24 Oct, 2024	112	WOOFi Swap	5 Mar, 2024
33	EasyFi	19 Apr, 2021	73	UwU Lend	10 Jun, 2024	113	Platypus Finance	16 Feb, 2023
34	Uranium Finance	28 Apr, 2021	74	Pickle	22 Nov, 2020	114	Moola Market	18 Oct, 2022
35	CoinEx	12 Sep, 2023	75	CREAM Lending	30 Aug, 2021	115	ALEX	6 Jun, 2025
36	Ooki	5 Nov, 2021	76	Snowdog	25 Nov, 2021	116	Gamma	21 Dec, 2021
37	Radiant V2	16 Oct, 2024	77	BearnFi	17 May, 2021	117	Thorchain DEX	22 Jul, 2021
38	Infini	24 Feb, 2025	78	Curio	25 Mar, 2024	118	Bittensor	2 Jul, 2024
39	KyberSwap Elastic	22 Nov, 2023	79	Indexed Finance	14 Oct, 2021	119	Origin Protocol	17 Nov, 2020
40	Cashio	23 Mar, 2022	80	Team Finance	27 Oct, 2022	120	HTX	25 Sep, 2023

ID	Name	Date	ID	Name	Date	ID	Name	Date
121	LCX	8 Jan, 2022	161	Thorchain DEX	15 Jul, 2021	201	JayPegs Automart	17 Sep, 2021
122	Multichain	10 Jul, 2021	162	Loopring	9 Jun, 2024	202	Orion Pools	2 Feb, 2023
123	Warp Protocol	18 Dec, 2020	163	1inch Network	7 Mar, 2025	203	Skyward Finance	2 Nov, 2022
124	Meter Bridge	6 Feb, 2022	164	Poly Network	2 Jul, 2023	204	Fortress Loans	8 May, 2022
125	CoinDash	18 Jul, 2017	165	OcelotDex	20 Sep, 2024	205	DEUS Finance	15 Mar, 2021
126	Rho Markets	19 Jul, 2024	166	Super Sushi Samurai	22 Mar, 2024	206	Banana Gun	19 Sep, 2024
127	BNS	1 Feb, 2022	167	DeltaPrime	11 Nov, 2024	207	Swaprum	19 May, 2023
128	KiloEx	14 Apr, 2025	168	Tapioca DAO	18 Oct, 2024	208	Arena SocialFi	7 Oct, 2023
129	Jimbos Protocol	28 May, 2023	169	DuelBits	13 Feb, 2024	209	SUNRAY-FINANCE	30 Oct, 2024
130	BurgerSwap	28 May, 2021	170	pNetwork	3 Nov, 2022	210	Remitano	15 Sep, 2023
131	Exactly	18 Aug, 2023	171	xToken	29 Aug, 2021	211	Turtle Dex	20 Mar, 2021
132	DAO Maker Vesting	12 Aug, 2021	172	Gamma	4 Jan, 2024	212	CoinSpot	9 Nov, 2023
133	Hundred Finance	15 Apr, 2023	173	Eleven Finance	22 Jun, 2021	213	Bunny	16 Jul, 2021
134	ValueDefi	14 Nov, 2020	174	Radiant V2	2 Jan, 2024	214	Trinity Wallet	12 Feb, 2020
135	Lodestar V0	10 Dec, 2022	175	Raydium AMM	16 Dec, 2022	215	TempleDAO	11 Oct, 2022
136	Velocore V2	2 Jun, 2024	176	ChainSwap	7 Nov, 2021	216	Fantasm Finance	9 Mar, 2022
137	Holograph	13 Jun, 2024	177	ALEX	16 May, 2024	217	CoinStats	12 Jul, 2024
138	Abracadabra Spell	30 Jan, 2024	178	DAO Maker Vesting	4 Sep, 2021	218	Platypus Finance	13 Oct, 2023
139	Alchemix	16 Jun, 2021	179	Ionic Protocol	4 Feb, 2025	219	Mobius Token	11 May, 2025
140	Seneca	28 Feb, 2024	180	Kokomo Finance	27 Mar, 2023	220	Unizen	8 Mar, 2024
141	DEUS Finance	5 May, 2023	181	Ola Finance	31 Mar, 2022	221	Onyx Protocol	1 Nov, 2023
142	Magnate Finance	25 Aug, 2023	182	SurgeBNB	17 Aug, 2021	222	ETHTrustFund	22 Jul, 2024
143	Belt Finance	29 May, 2021	183	Onyx V2	26 Sep, 2024	223	Zunami Protocol	13 Aug, 2023
144	Hundred Finance	15 Mar, 2022	184	DFX V2	10 Nov, 2022	224	Gym Network	8 Jun, 2022
145	Defrost	24 Dec, 2022	185	ForceBridge	6 Jun, 2025	225	Revest Finance	27 Mar, 2022
146	LendHub	16 Jan, 2023	186	UwU Lend	13 Jun, 2024	226	MangoFarmSol	7 Jan, 2024
147	Audius	25 Jul, 2022	187	dForce	10 Feb, 2023	227	pump.fun	16 May, 2024
148	DeltaPrime	16 Sep, 2024	188	Dragoma	8 Aug, 2022	228	Sirio Finance	1 Feb, 2025
149	Bondly	15 Jul, 2021	189	Nirvana V1	28 Jul, 2022	229	Dexible V2	20 Feb, 2023
150	Loopscale	26 Apr, 2025	190	Siren	3 Sep, 2021	230	Mirror	8 Oct, 2021
151	Inverse Finance Frontier	16 Jun, 2022	191	Squid Games	3 Nov, 2021	231	DODO AMM	9 Mar, 2021
152	Roll	14 Mar, 2021	192	EraLend	25 Jul, 2023	232	MM Finance Cronos	4 May, 2022
153	Agave	15 Mar, 2022	193	Vesper Finance	3 Nov, 2021	233	Ovix	28 Apr, 2023
154	Slope Wallet	2 Aug, 2022	194	Shido	29 Feb, 2024	234	Akropolis	12 Nov, 2020
155	Truflation	25 Sep, 2024	195	Bungee	16 Jan, 2024	235	GemPad	17 Dec, 2024
156	ZKSync	15 Apr, 2025	196	Raft	10 Nov, 2023	236	Dough Finance	12 Jul, 2024
157	ParaSpace Lending V1	17 Mar, 2023	197	CheeseBank	16 Nov, 2020	237	Omm	21 Jan, 2023
158	PolyNetwork	12 Aug, 2024	198	SushiSwap	10 Apr, 2023	238	Hope Finance	20 Feb, 2023
159	Ankr	2 Dec, 2022	199	Conic Finance	21 Jul, 2023	239	MERLIN DEX	25 Apr, 2023
160	Terra 2.0	31 Jul, 2024	200	Zabu Finance	12 Sep, 2021	240	XBridge	24 Apr, 2024

ID	Name	Date	ID	Name	Date	ID	Name	Date
241	Nexera	7 Aug, 2024	281	FeGex	29 Dec, 2024	321	Merlin	29 Jun, 2021
242	Dolomite	20 Mar, 2024	282	Palmswap	24 Jul, 2023	322	Uwerx	2 Aug, 2023
243	8ightDAO	8 Dec, 2021	283	RocketSwap Base	15 Aug, 2023	323	Locus Finance	30 Dec, 2023
244	Bent Finance	21 Dec, 2021	284	ElasticSwap	13 Dec, 2022	324	Bond Protocol	21 Oct, 2022
245	Concentric	22 Jan, 2024	285	Orange Finance	8 Jan, 2025	325	Rivus DAO	16 Sep, 2024
246	Grand Base	15 Apr, 2024	286	GMBL.COMPUTER	6 Sep, 2023	326	Amun	26 Dec, 2023
247	uniBTC	26 Sep, 2024	287	SudoRare	23 Aug, 2022	327	MoonHacker	2 Jan, 2025
248	Paraluni Masterchef	13 Mar, 2022	288	Balancer V2	27 Aug, 2023	328	Impermax V3	26 Apr, 2025
249	BonqDAO	1 Feb, 2023	289	Sturdy V1	12 Jun, 2023	329	Conic Finance	22 Jul, 2023
250	Tender Finance	7 Mar, 2023	290	AutoShark	24 May, 2021	330	KyberSwap Classic	1 Sep, 2022
251	Acala Network	13 Aug, 2021	291	1inch Network	31 Oct, 2024	331	LeadBlock	14 Oct, 2024
252	Pike	30 Apr, 2024	292	LaunchZone	27 Feb, 2023	332	Tectonic	22 Feb, 2024
253	Levyathan	30 Jul, 2021	293	Merlin	26 May, 2021	333	Safe Dollar	28 Jun, 2021
254	YOLO Games	11 Jun, 2024	294	Midas Capital	17 Jan, 2023	334	Balancer V2	19 Sep, 2023
255	Moby	8 Jan, 2025	295	Hector Lending	2 Jun, 2023	335	Astrid Finance	29 Oct, 2023
256	Florence Finance	30 Nov, 2023	296	Allbridge Core	2 Apr, 2023	336	Convergence	1 Aug, 2024
257	Rubic	25 Dec, 2022	297	Ooki	18 Feb, 2020	337	zkFinance	11 Nov, 2024
258	Solareum	29 Mar, 2024	298	Unibot	31 Oct, 2023	338	Spectral Labs	1 Dec, 2024
259	Treasure	3 Mar, 2022	299	LeetSwap	1 Aug, 2023	339	Keep3r Network	12 Jun, 2023
260	Minterest	15 Jul, 2024	300	Midas Capital	17 Jun, 2023	340	Unilend V2	12 Jan, 2025
261	BlueBerry	24 Feb, 2024	301	Blur Finance	10 Aug, 2022	341	four.meme	11 Feb, 2025
262	Growth DeFi	9 Feb, 2021	302	Curve DEX	9 Aug, 2022	342	Yield Protocol	30 Apr, 2024
263	Bitcoin Mission	22 Apr, 2025	303	GMX V1	18 Sep, 2022	343	SharedStake	1 Sep, 2023
264	Monoswap	31 Jul, 2024	304	Merlin	26 May, 2021	344	Tropykus RSK	14 Jun, 2023
265	RiskOnBlast	25 Feb, 2024	305	Earning.Farm	9 Aug, 2023	345	Ara Finance	19 Jun, 2023
266	LND	9 May, 2025	306	Hedera	10 Mar, 2023	346	Lava	4 Oct, 2024
267	GeniusAI	5 May, 2024	307	Zunami Protocol	15 May, 2023	347	Magpie Protocol	23 Apr, 2024
268	ApeRocket	14 Jul, 2021	308	Clober Liquidity Vault	10 Dec, 2024	348	PeopleDAO	8 Mar, 2023
269	Save	2 Nov, 2022	309	Vestra DAO	4 Dec, 2024	349	Local Traders	24 May, 2023
270	Lendf.me	19 Apr, 2020	310	Maestro	23 Oct, 2023	350	Atlantis Loans	10 Jun, 2023
271	Steadefi	8 Aug, 2023	311	Wise Lending V1	12 Jan, 2024	351	WXETA	16 Sep, 2024
272	Kannagi	30 Jul, 2023	312	Arcadia V1	10 Jul, 2023	352	ArbiSwap	2 Mar, 2023
273	Levana Perps	27 Dec, 2022	313	Clipper	1 Dec, 2024	353	Hegic(old contract)	1 Mar, 2020
274	Sovryn	4 Oct, 2022	314	Cardex	18 Feb, 2025	354	Tectonic	18 Nov, 2024
275	Cypher	7 Aug, 2023	315	Themis Protocol	27 Jun, 2023	355	Rodeo	4 Jul, 2023
276	Level Perps	1 May, 2023	316	SIR	30 Mar, 2025	356	BTC24H	16 Dec, 2024
277	NowSwap	15 Sep, 2021	317	Ooki	15 Feb, 2024	357	Roe Finance	11 Jan, 2023
278	DeltaPrime	23 Jul, 2024	318	Kalax	14 Oct, 2024	358	AzukiDAO	3 Jul, 2023
279	Sentiment	4 Apr, 2023	319	Zircon Gamma	18 Mar, 2023	359	Ratio Finance	3 Dec, 2022
280	OKX NFT Aggregator	20 Jun, 2024	320	Lava	29 Mar, 2024	360	Hyperliquid	15 Jun, 2023