5-2023

The role of disclosure in DeFi markets: Evidence from Twitter

Amanda AW YONG

Singapore Management University

THE ROLE OF DISCLOSURE IN DEFI MARKETS: EVIDENCE FROM TWITTER

AMANDA AW YONG

SINGAPORE MANAGEMENT UNIVERSITY

2023

The Role of Disclosure in DeFi Markets: Evidence from Twitter

Amanda Aw Yong

Submitted to School of Accountancy

In partial fulfillment of the requirements for the

Degree of Doctor of Philosophy in Accounting

Dissertation Committee:

Liandong Zhang (Chair)
Lee Kong Chian Professor of Accounting
Singapore Management University

Richard Crowley
Assistant Professor of Accounting
Singapore Management University

Weikai Li
Assistant Professor of Finance
Singapore Management University

Holly Yang
Associate Professor of Accounting
Singapore Management University

Singapore Management University

2023

I hereby declare that this PhD dissertation is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree in any university previously.

Amanda Aw Yong

30 April 2023

ABSTRACT

Decentralized Finance (DeFi) platforms use self-executing smart contracts to provide financial services and are programmed to automatically post all transactions on the public blockchain. Notwithstanding this public availability of blockchain information, DeFi platforms also aggregate these transactions and disclose the summarized blockchain information on their Twitter accounts. This paper studies whether and how voluntary disclosure of blockchain information plays a role in the transparent DeFi market. I find that the number of blockchain-related tweets is associated both with an increase in the platform's Total Value Locked (TVL) and with an increase in the unique number of platform users. The relationship between blockchain-related tweets and TVL is strengthened when the tweets have greater information content and when users face higher information processing costs. This suggests that public blockchain transactions are too costly for users to process such that they rely on the platform's disclosure. Overall, my results show that DeFi platforms can help users process and understand blockchain transactions by summarizing and disclosing them on Twitter.

JEL codes: G10, G24, M40, M41

Keywords: decentralized finance (defi), cryptocurrency, disclosure, information processing costs

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

How do you eat an elephant? One bite at a time. This dissertation is but a pit-stop in my quest for an academic elephant and owed to many.

To my advisor, Liandong, I am forever indebted. For inspiring me to think deeply and clearly, for nudging me in the right direction with thoughtful questions, and for infecting me with your passion for research. Above all, thank you for knowing just the right thing to say when the going gets tough and for challenging me to go further when I get complacent. This dissertation, if not this PhD journey, would be insurmountable without your guidance.

To my mentor, Richard, thank you for the cup of coffee at Kurasu 5 years ago. I asked you then why there was no Fintech research using Twitter data. Throughout the years, your mentorship over programming and data finally culminated in this dissertation. Thank you for believing in me every time I was drowning in self-doubt and then teaching me how to swim to shore.

To my committee member, Holly, thank you for showing and teaching me what academic research is from the very first PhD seminar in year 1. The papers and lively discussions have greatly shaped and influenced my mental schema. For always being available to chat and for providing academic and life advice, my PhD journey has been made more meaningful.

To my committee member, Weikai, I have learned from you about the importance of doing research that speaks to the heart. From our PhD seminar to our interactions, your research ethic has and will continue to inspire me.

A large part of the academic journey is also owed to hallway chats and advice from faculty. To the faculty at SOA, I thank you for your empathy and concern for struggling PhD students like me, and for your constant encouragement that the dissertation will eventually be completed. To the friends I made on this journey and especially my batch mates Tian and Huiyu, the random office visits and writing equations on our whiteboards are memories that I keep close to my heart whenever this journey feels lonely.

To the TRC girls, Dreams Come True, Sweej and Drea who waited patiently while I was in the deep academic abyss and then buying me dinner whenever I emerged, thank you for providing emotional and physical fuel to get me through the past 5 years. To my family who supported my decision to embark on this academic journey and then supported me in every possible way, your love and sacrifices are deeply felt.

“How do you eat an elephant? One bite at a time.” To my husband, James, who walked beside and carried me every step of this quest, thank you for your elephant wisdom on the darkest nights and for celebrating every milestone big or small. For not only being my biggest cheerleader, but also challenging every idea I ever had and for pushing me to think creatively. To my toddler Joy who have taught me more about life than I have taught you, I hope one day you find your elephant quest.

Above all, glory to God, who is able, through his mighty power at work within us, to accomplish infinitely more than we might ask or think.

DEDICATION

For James and Joy

1. Introduction

Decentralized Finance (DeFi) is a general term for an evolving trend in the cryptocurrency ecosystem – a group of blockchain-based decentralized applications providing financial services such as lending, cryptocurrency exchange, and assets management (Gogel 2021). Unlike other financial service providers, DeFi platforms do not act as centralized intermediaries but instead rely on self-executing smart contracts that algorithmically facilitate transactions. DeFi users may directly monitor the execution of these transactions as both the smart contract code and transaction details are publicly observable from the blockchain in real time.[1] Such unprecedented transparency and innovation resulted in the immense popularity of DeFi platforms. Total Value Locked (TVL)[2] in DeFi platforms grew from US$9.1 billion in June 2020 to US$86.6 billion in February 2022 (DeFiPulse 2022). However, this exponential growth is marred by substantial financial losses. Chainalysis estimates that $2.2 billion was stolen from DeFi platforms in 2021, most of which exploited errors in the smart contract code that these platforms operate on (Chainalysis 2022). It thus appears that DeFi users are unable to accurately assess platform credibility, causing them to deposit funds in platforms with coding vulnerabilities.

[1] For example, users can rely on blockchain explorers such as Etherscan to view both pending and completed transactions. Transactions are recorded on the blockchain when they have been completed and details such as the transaction amount, transaction fee, transaction ID, counterparty details, and timestamp, are publicly observable. In addition, blockchain explorers also publish the smart contract code that defines transaction terms.

[2] TVL is the overall value of cryptocurrency assets deposited in DeFi platforms by DeFi users and is a key metric for measuring user confidence in the platform (https://www.coindesk.com/learn/why-tvl-matters-in-defi-total-value-locked-explained/). For a more detailed explanation of TVL, please refer to Appendix A.

This difficulty in assessing platform credibility has increasingly been a focus of regulators' attention. The Financial Stability Board (FSB) highlights the "lack of…reporting requirements producing consistent and reliable data" as an important factor impeding users' understanding of DeFi platforms (Financial Stability Board 2023). Similarly, SEC Commissioner Caroline Crenshaw asserts that "absent mandatory disclosure requirements, information asymmetries will likely disadvantage…the smallest investors", preventing them from "assess[ing] risk likelihood and severity" of dealing with DeFi platforms (Crenshaw 2021). The key risk in transacting with lending platforms is that they give out too many loans such that borrowers default and lenders cannot withdraw their assets (Financial Stability Board 2023). Bertomeu et al. (2022) shows that this risk can be estimated by aggregating all borrowing and lending transactions from the public blockchain. Given that DeFi users can directly obtain information to estimate such risks, regulators' focus on a lack of disclosure requirements prompts an interesting question: is disclosure necessary in a transparent DeFi market?

In this paper, I examine whether and how voluntary disclosure[3] of blockchain-related information by DeFi platforms affects their TVL, where TVL is measured as the total value of cryptocurrency assets that users deposit in DeFi platforms in exchange for financial services. The DeFi market uses TVL as a key performance indicator as it represents the financial commitment that

[3] Some studies use "dissemination" specifically for the same information already disclosed elsewhere (e.g., Blankespoor et al. 2014), while many others use "disclosure" even if the information disclosed elsewhere is largely similar (e.g., Christensen et al. 2017). To be consistent with the majority of accounting literature, I use the term "disclosure" with a broad definition to refer to both disclosure of summarized blockchain information (e.g., interest rate aggregated from blockchain transactions) and dissemination of identical information that can be found on the blockchain (e.g., specific transactions).

DeFi users make to the platform. Despite a lack of disclosure regulation, DeFi platform founders often aggregate and summarize blockchain transactions before disclosing them on Twitter. Such voluntary disclosure is likely driven by economic incentives to help users process information as the founders' stake becomes more valuable when their platforms have more users and higher TVL (Cong et al. 2021).

On the one hand, voluntary disclosure of blockchain information may be redundant in a unique setting like DeFi where platforms are programmed to be transparent and the information "stored on the underlying blockchain can be publicly scrutinized" (Schär 2021). Besides analyzing blockchain transactions to estimate borrower default risk as demonstrated in Bertomeu et al. (2022), DeFi users can also directly examine the smart contracts that execute DeFi platform transactions to ascertain if there are any coding vulnerabilities and timely code upgrades. With the public availability of blockchain information, DeFi users could directly assess if the platform is credible without relying on blockchain-related disclosure. On the other hand, blockchain-related disclosure may be informative as the high costs of processing blockchain data could cause users to ignore publicly available information (Blankespoor et al. 2020; Hu et al. 2022). The granular transaction-level data that is posted on the blockchain, while useful, requires high computing power and advanced programming skills that may not be accessible for all DeFi users. Crenshaw asserts that "it is not reasonable to build a financial system that demands investors also be sophisticated interpreters of complex code" (Crenshaw 2021) and the FSB underscores "the difficulty in aggregating…[the] vast amount of data available on [the blockchain]" (Financial Stability Board 2023). Just as how capacityconstrained investors prefer summarized financial information such as total sales over disaggregated information such as multiple sales components (Lu 2022), DeFi users facing high processing costs could rely on summarized blockchain-related disclosure by DeFi platforms and ignore detailed blockchain data.

To examine whether voluntary disclosure of blockchain information affects DeFi platforms' TVL, I first construct a sample of 131 DeFi platforms that provide a wide range of services such as decentralized exchanges (DEXes), derivatives, payments, assets management, and lending. I then hand-collect their financial information from DeFi Pulse and Coinmarketcap and their social media disclosures from Twitter, for the period January 2018 to March 2022. Of the 131 platforms, 121 platforms have an active Twitter account, suggesting that Twitter is a key communication channel for DeFi platforms. This parallels prior accounting research that provides evidence of Twitter being an effective disclosure channel for S&P1500 firms (Blankespoor et al. 2014; Crowley et al. 2022). As DeFi platforms also use Twitter for other purposes such as advertising and customer support, I construct my disclosure variable by identifying and including only blockchain-related tweets through a machine-learning method as described in Section 3. My final sample comprises 106 unique platforms and 2,150 platform-month observations with non-missing variables.

Next, I test the association between the number of blockchain-related tweets and TVL. Since TVL is a measure of users' confidence in DeFi platforms (George 2022), a positive association suggests that voluntary disclosure of blockchain information is helpful for DeFi users to assess credibility such that users increase their confidence in, and financial commitment to, the platforms.

Accordingly, I find that the number of blockchain-related tweets in the prior period is positively associated with TVL in the current period. My results hold after controlling for users' demand for blockchain information and platform and year fixed effects. Additional tests show that when broken down by blockchain topics, users react to disclosure about current yield, platform governance, and platform features.

If blockchain-related disclosure is informative for DeFi users, then tweets with greater information content should be more positively associated with the platform's TVL. I first follow the intuition in Dyer et al. (2017) that repeated sentences across documents can be redundant and less informative. To calculate the extent of content similarity between tweets, I employ a machine learning technique, Universal Sentence Encoder, developed by Google (Cer et al. 2018) and recently used in the accounting literature (Crowley et al. 2021). The second textual characteristic I examine is the number of topics that platforms tweet about. The more blockchain-related topics that platforms discuss, the greater the information content of their tweets. Consistent with expectations, cross-sectional tests show that the sub-samples with greater content difference and higher number of topics demonstrate significantly larger associations between the number of blockchain-related tweets and TVL.

Furthermore, I provide evidence for my argument that disclosure matters due to high information processing costs that users face. As platforms are operated by smart contracts, I first calculate the lines of smart contract code[4]

[4] Lines of code is a common measure used in computer science to measure the complexity of the source code. The more lines of code used, the more complex the source code is (Albrecht and Gaffney 1983).

that each platform uses. This variable is both a direct and indirect measure of information processing costs. For more technical users, it directly captures the time and effort spent to understand and evaluate the smart contract code. For more general users, the lines of code indirectly capture the complexity of the platform structure. Thus, platforms using more lines of contract code have higher processing costs and users should reap more benefits from disclosure. In addition, I count the lines of code that DeFi Pulse uses to calculate platforms' TVL. More lines of code correspond with the difficulty of extracting smart contract information from the platform. Consistent with expectations, I find that my results are driven by the sub-samples with more lines of contract code and TVL code. Taken together, these results support my argument that disclosure of blockchain information is helpful for users because they face high costs of processing public blockchain data.

In addition, I use an alternative dependent variable to alleviate the concern that TVL may fluctuate along with ETH price and not always reflect true platform activity (see Appendix A). DeFi users use their cryptocurrency wallets to transact with platforms and these wallets are identifiable by a unique wallet address (conceptually like bank account numbers). I thus use the number of unique addresses to proxy for the number of users and hypothesize that disclosure could help more users process blockchain data. Accordingly, I find that the number of blockchain-related tweets is associated with a greater number of platform users and when broken down by user type, is associated with both new and returning users. I also find that disclosure does not increase TVL per user, which suggests that disclosure increases TVL by increasing the total number of users (i.e., extensive growth).

I next explore if greater salience about the importance of transparency affects the relationship between disclosure and TVL. Unlike DeFi platforms, centralized cryptocurrency exchanges (CEXes) are opaque and their transactions are not posted on the public blockchain. Thus, when CEXes experience cybersecurity hacks, users are unable to monitor transaction activity and cannot directly gather information about how the hack occurred or whether their cryptocurrency assets are intact. Adverse events such as hacks thus increase the cryptocurrency community's attention to the importance of transparency. I find that when CEXes suffer a hack, platforms that disclose more blockchain-related information attract a greater number of users and higher TVL. This strengthens my inference that disclosure is informative for DeFi users especially when they are primed to focus on transparency.

My study is subject to several caveats. First, there is a possibility of reverse causality, where it is TVL that is driving the number of blockchain-related tweets. This concern is less severe as I show that TVL in the prior period does not significantly affect blockchain-related tweets (see Appendix D) and I employ a lead-lag structure in all my analysis by regressing the lagged value of disclosure on the lead value of TVL. My results also hold after controlling for TVL in the prior period. Nevertheless, I conduct an additional test using Google searches of "Ethereum" to proxy for user interest as platforms may tweet more when there is greater interest in the underlying Ethereum technology. I find that my results are significant in both periods of high and low user interest and that the difference between the subsamples is statistically indistinguishable. Notwithstanding these mitigation strategies, my results are still meaningful in the presence of such concerns. If TVL is indeed driving the number of tweets, then it suggests that platforms react to DeFi users' demand for information and thus increase their disclosure on Twitter. Thus, both disclosure and Twitter as a disclosure channel remain important.

Second, there may be correlated omitted variables that affect both TVL and blockchain-related tweets. Besides controlling for Ether (ETH)'s price momentum and including year fixed effects to hold constant macro-level factors affecting DeFi markets, I also include platform fixed effects to reduce concerns that my results are driven by platform-specific factors and further exploit the textual content of tweets to show that my results vary with information content of tweets. Furthermore, my results are similar when regressing TVL on tweets at the daily level and when regressing change in TVL on change in tweets, both of which are ostensibly less subject to omitted variables. Finally, I follow the method developed by Frank (2000) to show that there is only a small likelihood that a confounding variable renders my main result insignificant. Although these tests mitigate empirical concerns, I acknowledge that my results rely on association-based tests and thus refrain from drawing causal inferences.

Third, there might be an alternative explanation where blockchain-related tweets increase TVL due to an advertising effect instead of an information effect as I hypothesize. There are two types of advertising tweets in my setting. The first is a pure advertising tweet: "Kyber Network is currently the most popular exchange to trade DAI on!". These tweets are not included in my sample as they do not contain blockchain information. The second type is included in my sample as they are blockchain-related, but they may also be construed as advertising: "Bancor has now an $SNX liquidity pool!". This new feature is coded into the smart contract and users would be aware if they had been tracking changes to the smart contract code. To rule out a pure advertising effect, I further restrict my sample to tweets with concurrent changes to the smart contract code as advertising campaigns would not necessitate code changes. I show that my results are driven by the sub-sample where the platform also made changes to their smart contract code, which suggests that the tweets were discussing new blockchain information. This provides support that blockchain-related tweets increase TVL due to an information effect.

My paper first provides practical implications for the DeFi market that is growing in prominence within the cryptocurrency ecosystem. Since the collapse of FTX, one of the largest centralized cryptocurrency exchanges, the cryptocurrency community has been espousing on the virtues of decentralization. Ethereum blockchain co-founder Vitalik Buterin opined that FTX's collapse led many to see that "centralized anything is by default suspect" (Shukla et al. 2022). While the draw of DeFi is that users retain control over their assets and have full visibility of platform transactions, the reality is that blockchain transactions are inherently challenging to process and understand. My paper provides evidence that platform founders can address this challenge by summarizing and disclosing blockchain transactions on Twitter. In turn, such voluntary disclosure increases the platform's TVL and contributes to the overall financial health of DeFi markets.

In addition, my paper could provide insights to regulators who have been debating over how to regulate the cryptocurrency ecosystem. President Biden signed an Executive Order in March 2022 calling for a comprehensive understanding of blockchain and DeFi to foster responsible development of digital assets. Subsequently in February 2023, SEC Chair Gary Gensler asserted that cryptocurrency service providers must "provide full, fair and truthful disclosure" (SEC 2023). While the notion of regulation seems like an eventuality, how disclosure regulation should be designed seems less straightforward. My results show that the summary of blockchain information is helpful for users and that users value disclosure content such as current yield, platform governance, and platform features.

I next contribute to the emerging literature on blockchain technology and DeFi by being the first to document voluntary disclosure practices within the DeFi market. Two recent papers are most closely related to my work. First, Bourveau et al. (2022) study initial coin offerings (ICO) whitepapers and find that the greater the disclosure, the better the ability to raise capital. Information disclosed in ICOs are likely non-blockchain data such as roadmap for product development and expected use of proceeds (Bourveau et al. 2022). In the DeFi setting, however, I study the disclosure of blockchain information that is likely observable from the blockchain. Second, Hu et al. (2022) examines trader learning within a specific DeFi lending platform. They find that traders ignore public information and do not mimic well-performing traders. While Hu et al. (2022) examines how processing costs prevent individuals from mimicking profitable strategies, my paper shows how disclosure of blockchain information increases the platform's TVL due to information processing costs. My paper is thus a first attempt at understanding the information environment within the DeFi market.

More broadly, my results speak to the information processing literature by extending beyond the stock market setting to an alternative financial ecosystem with unique features. Prior studies have examined how information processing costs prevent investors from effectively using public disclosure (Blankespoor et al. 2020). For example, Christensen et al. (2017) shows that republishing safety violations in SEC filings has positive real effects as stakeholders may have been unaware of the same information available on the regulator’s website. Thus, Christensen et al. (2017) focuses on how the SEC filing dissemination channel increases awareness of information. In contrast, I focus on how summarized disclosure of disaggregated data helps users process information in the new and emerging DeFi market. Its unique feature is that granular blockchain transactions are publicly available, which allows me to estimate the effect of summarizing disaggregated information as opposed to the effects of disclosing summary information not available elsewhere. My paper thus highlights the positive role of voluntary disclosure for platforms and users in the DeFi market due to information processing costs.

2. Institutional Background

2.1 Key Features of DeFi Platforms

DeFi is a category of blockchain-based decentralized applications (Dapps) for financial services (Deshmukh et al. 2021). There are two distinct characteristics of DeFi platforms. First, the operational system of platforms can be publicly scrutinized. DeFi platforms are fully operated by ‘smart contracts’, which is a set of code used to represent and execute contractual agreements (Zetzsche et al. 2020). These agreements are programmatically executed once the pre-agreed upon terms have been met and this feature minimizes the risk of manipulation and arbitrary intervention (Schär 2021). For example, if a DeFi user A agrees to lend 10 Bitcoins to another DeFi user B, they may both agree for B to over-collateralize and lock in collateral of 5,170 USDC (a cryptocurrency backed by fiat money, i.e., stablecoin). However, if the value of 5,170 USDC later falls below the agreed upon collateral value due to exchange rate fluctuations, the smart contract will automatically liquidate the collateral to facilitate loan repayment to A.[5] In order to change a smart contract feature (e.g., the formula used to calculate interest rate), platforms typically conduct governance voting proposals on the public blockchain. DeFi users who hold platform tokens are eligible to vote on these proposals. Platform tokens are initially distributed by the platform to users who engage the platform's financial services and tokens could later be listed and traded on cryptocurrency exchanges. Thus, since DeFi platforms are fully operated by smart contracts and the corresponding smart contract code is publicly observable, DeFi users have full access to the platforms' operational system.

Second, all transaction activities of DeFi platforms are publicly available on a fully disaggregated basis. The smart contracts for DeFi platforms are executed on a public, permissionless blockchain (Chen and Bellavitis 2020) that allows anyone and everyone full access to information stored on-chain (Buterin 2015). Most DeFi platforms operate on the Ethereum blockchain, which allows platforms to execute smart contracts that its predecessor, Bitcoin, could not (Cointelegraph 2022). A permissionless blockchain executes smart contracts by requiring a large group of validators to verify that the transactions

[5] Many DeFi lending platforms require borrowers to over-collateralize to provide assurance to lenders as there is no borrower screening as in traditional financial markets. For example, Aave (a DeFi lending platform) requires borrowers to over-collateralize at 116% their loan amount when they use USDC as collateral (https://docs.aave.com/risk/asset-risk/risk-parameters). Thus, to borrow 10 BTC, borrowers must lock in USDC amounting to 11.6 BTC (~5,170 USDC). Aave allows a slight buffer of 3% for BTC/USDC exchange rate fluctuations. If the exchange rate of BTC/USDC changes such that the value of 5,170 USDC is worth less than 113% of BTC, then the smart contract goes into an automatic liquidation process. This is where any user can act as a liquidator and repay the BTC loan on behalf of the borrower and earn a liquidation fee from the USDC collateral (https://docs.aave.com/faq/liquidations).

adhere to the agreed upon terms. During my sample period, Ethereum operates on a Proof-of-Work system where validators (also known as miners) compete for the right to verify transactions by solving a computationally intensive problem (Makarov and Schoar 2022).[6] Once the validators solve the computation and agree that the transaction is based on contractual terms, the transaction is successfully added to the blockchain and validators receive an economic reward for their work (Makarov and Schoar 2022). Both the transactions pending validation and verified transactions are publicly observable from the blockchain. Users typically use blockchain explorers such as Etherscan to view these transactions and they can observe details such as transaction hash (i.e., unique transaction ID), wallet addresses of contracting parties (conceptually similar to bank account numbers), transaction cost (paid to validators), timestamp, and transaction value. Thus, since blockchain information is available at the transaction level, users have access to disaggregated information about the platforms.

Taken together, these characteristics create a new financial architecture that does not require human intermediaries and as a result are completely transparent (Makarov and Schoar 2022). DeFi users may directly access the operational system of the platform by scrutinizing its smart contract code. This enables them to identify potential operational weaknesses (e.g., coding errors) and to decide if they agree to the transaction terms defined in the code. If such evaluation is positive, users may then deposit their cryptocurrency assets and track their intended transactions being executed in real-time. DeFi platforms

[6] Ethereum transited to a Proof-of-Stake system on 15 September 2022, which is out of my sample period.

thus provide a unique research setting where both the operational system and transaction activities are completely transparent and publicly accessible.

2.2 Economic Incentives of DeFi Platform Founders

DeFi platforms are operated by smart contracts written by platform founders. Besides programming platforms to provide financial services, most platform founders also create platform tokens that typically carry voting rights. Tokens may be given to platform users (i.e., what the DeFi community term as "Airdrops") or bought from cryptocurrency exchanges from other users. Token holders can then vote on changes to platform features such as accepting a new cryptocurrency as loan collateral or the methodology to calculate interest rates. Thus, the more users and funds a platform has, the more valuable the tokens with voting rights (Cong et al. 2021).

These tokens serve as an important economic incentive for platform founders. Founders typically retain a portion of the tokens upon creation and they may "cash out" their invention by selling the tokens. With this vested interest, founders are then incentivized to increase the demand for its platform's services as high platform usage drives future token appreciation (Cong et al. 2021). One way through which founders can increase platform adoption is to facilitate users' assessment of platform credibility. For example, certain asset management platforms may have been consistently earning high returns for their users. While users can calculate these historical returns from blockchain transactions, handling such voluminous transactions requires high computing power and advanced programming skills. Platform founders could instead process these transactions and disclose the summarized returns to its users.

Doing so increases user confidence that the platform can deliver good financial performance and this in turn attracts more users and funds to the platform. Thus, platform founders typically process and disclose summarized blockchain information due to their economic incentives.[7]

2.3 Blockchain-Related Information and Disclosure

DeFi platforms are unregulated and do not have human intermediaries facilitating transactions. There is thus strong advocacy within the DeFi community to "do your own research" (DYOR) which SEC Commissioner Crenshaw describes as the "buyer beware" approach (Crenshaw 2021). This means that DeFi users should obtain information to assess platform credibility as they completely assume the risk of any and all losses from transacting with DeFi platforms. The feasibility of DYOR rests on the premise that both the smart contract code and detailed blockchain transactions are publicly observable, rendering it possible for users to scrutinize and accurately evaluate DeFi platforms (Schär 2021). For example, a key risk of lending platforms is that they have insufficient liquidity when lenders want to withdraw their funds. Bertomeu et al. (2022) shows that the risk of lending platforms can be calculated by obtaining detailed information about borrowing and lending activities from the Ethereum blockchain. Similarly, DeFi users can compute these risk measures for the different platforms that they intend to transact with before

[7] A natural question that arises is whether founders disclose good news and hide bad news. There are two challenges in answering this question. First, the blockchain information that platforms disclose are mostly qualitative in nature. This prohibits a clear definition of good vs. bad news, analogous to how we compare actual vs. forecasted EPS in traditional capital markets. Second, as DeFi markets are currently unregulated, platform founders are not obliged to disclose bad news. From my analysis of the data, I indeed find "negative news" disclosure so rare that it may not provide sufficient statistical power. Thus, this paper is premised upon founders' economic incentives to help users process and understand blockchain information, instead of exploring strategic disclosure incentives.

deciding on the most credible platform. However, doing so is challenging because each lending platform is structured differently, and users must first figure out how the total deposits and loans are reflected in the different smart contracts that are linked to the platform. After which, users would need programming skills to extract the borrowing and lending activities linked to the relevant smart contracts and then possess sufficient computing power to handle the voluminous blockchain transactions. Evidently, relying on granular blockchain transactions makes DYOR a challenging task.

An alternative to aggregating blockchain transactions is to rely on summarized disclosure by DeFi platforms. Despite the lack of disclosure regulation, many platforms disclose summarized blockchain-related information through the platform's official Twitter account as founders have economic incentives to increase platform usage. I use an unsupervised machine learning method (i.e., I do not pre-determine any keyword list or potential topics) to provide insights on the type of blockchain information that platforms typically disclose (see Appendix B for a breakdown of topics). First, information regarding blockchain transactions helps users assess how trusted that platform is. Frequent transactions coupled with a high number of unique wallet addresses depositing cryptocurrencies suggest that many other users have confidence in the platform and thus lock their assets there. In addition, financial metrics such as aggregated transaction levels and user volume could be especially helpful for potential users choosing between different platforms to deposit their assets into. Second, information about governance voting proposals describes possible platform changes that users can evaluate. Proposals are typically conducted on the blockchain and users can monitor the platform's blockchain to stay updated on new voting proposals and to observe the voting results. Posting about these voting proposals on Twitter can reduce users' need to constantly monitor the blockchain for new proposals. Furthermore, disclosing the platform features that have changed as a result of governance proposals could also be informative for users. Third, disclosing the summarized current yield and past financial metrics could give users confidence that the platform is performing as well as they have previously promised to (i.e., that users are indeed enjoying the high yields as promised).

2.4 Hypotheses Development

This paper examines whether and how voluntary disclosure of blockchain information by DeFi platforms affects their TVL. Ex ante, it is unclear if disclosure of public information would result in any capital market benefits. The lack of regulation and advocacy for DYOR may incentivize DeFi users to directly process blockchain transactions such that the summarized disclosure does not provide users with additional information. Furthermore, prior accounting research finds that a greater degree of disaggregation represents higher disclosure quality (Chen et al. 2015). DeFi users may find the disaggregated blockchain transactions more informative than summarized blockchain information on Twitter and thus do not react to the latter.

However, the high cost of processing blockchain transactions may cause users to ignore public information (Blankespoor et al. 2020). Blockchain data is extremely granular and provided at the transaction level, thus requiring users to possess high computing power and advanced programming skills. In the presence of high processing costs and limited processing capacity, users may be unable to fully process disaggregated information. Lu (2022) analytically shows that investors presented with multiple detailed information have to optimally allocate their limited capacity between different components and are unable to fully process each piece of information. This results in them being better off with a single piece of summary information (e.g., total cost of sales) as compared to multiple components (e.g., line items adding up to cost of sales) even though the latter theoretically provides more information. Empirical studies have also shown that financial statement disaggregation increases processing costs thus necessitating greater auditor effort and audit fees (Beck et al. 2022; Koh et al. 2022). Similarly in my setting, the disaggregated nature of blockchain transactions may be too costly for users to fully process. DeFi users may instead only process the summarized information that platforms post on Twitter. As a result, voluntary disclosure helps users process public blockchain information to assess platform credibility and thus results in capital market benefits such as an increase in TVL. Thus, I state my first hypothesis:

H1: The number of blockchain-related tweets that DeFi platforms post on Twitter is positively associated with TVL.

One possible explanation why voluntary disclosure of public blockchain information may increase TVL is the presence of information processing costs. Users may find it too costly to directly process disaggregated blockchain transactions especially if the platform's structure is complex and difficult to understand or if users are new to the platform and face higher startup cost in understanding how the platform operates. Thus, an alternative for users facing high processing costs is to rely instead on aggregated blockchain information posted on Twitter as it is more easily understood. Thus, I state my second hypothesis:

H2: The number of blockchain-related tweets is positively associated with TVL because users face high information processing costs.

3. Research Design

3.1 Sample and Data

To construct my sample, I first obtain the names of all DeFi platforms listed on DeFi Pulse. DeFi Pulse is a data aggregator for DeFi platforms and it curates a list of platforms that must meet four criteria. First, platforms must be available on the Ethereum blockchain, as Ethereum is the optimal standard for smart contracts from a security and maturity perspective. Second, the smart contracts that execute platform services must be truly decentralized. Third, the platform's smart contract code must have undergone an audit. Finally, the smart contract must have safety features in place to prevent code administrators from exploiting other users (DeFiPulse 2021). At the time of download, there were 131 DeFi platforms listed on DeFi Pulse. From this list, I then searched for the platforms' Twitter pages through links provided on DeFi Pulse and supplement them with manual searches. This resulted in 121 platforms with active Twitter pages. After requiring non-missing control variables, I arrive at my main sample of 106 platforms, which provide a wide range of services such as DEXes, derivatives, payments, assets management and lending. The earliest observation of these 106 platforms with non-missing variables is in May 2018 and I conclude my data download in March 2022. Thus, my final sample comprises a platform-month sample of 2,150 observations, as reported in Table 1 Panels A and B.

3.2 Key Variables

My main outcome variable is TVL,[8] which is denominated in USD and provided by DeFi Pulse. I calculate the monthly average of TVL for each platform and use the natural logarithm of TVL in my main analysis (*Ln(TVL)*). In addition, I obtain from DeFi Pulse the monthly unique number of wallet addresses transacting with each platform to proxy for the number of users (*Ln(Unique Users)*). DeFi Pulse further decomposes unique users into new users (*Ln(New Users)*), returning users (*Ln(Returning Users)*), and resurrected users (*Ln(Resurrected Users)*). Returning users are defined as wallet addresses that have interacted with the platform in the previous month, while resurrected users are those that have previously interacted with the platform but did not do so in the previous month. As the data on platform users are provided on a monthly basis, I conduct all of my analysis at the monthly level in this paper.

Next, to measure voluntary disclosure, I first downloaded all tweets posted by DeFi platforms within my sample period. As platforms also use Twitter for other purposes such as advertising and customer support, it is important to restrict my analysis to tweets that contain blockchain-related information. To identify tweets with information content about the blockchain, I first use an unsupervised machine learning technique (Twitter-LDA) to categorize all tweets into 15 topics.[9] Next, I randomly select 100 tweets from each topic (i.e., total of 1,500 tweets) and assign a blockchain-related label if the content is related to the blockchain. For each topic, if the number of

---

[8] Appendix A presents a more detailed explanation of how TVL is calculated and the limitations of using TVL.
[9] Appendix B1 lists all 15 LDA topics together with top keywords from each topic.

blockchain-related labels exceed 80%, then all tweets belonging to that topic will be classified as blockchain-related tweets.[10] My main disclosure variable is the natural logarithm of the total number of tweets containing blockchain-related information (*Ln(Blockchain Tweets)*).[11]

I first control for overall cryptocurrency market sentiment (*ETH Momentum*) as my sample comprises of DeFi platforms that operate on the Ethereum blockchain and are thus affected by changes in the value of Ether (ETH), which is the token native to the Ethereum blockchain. In addition, I control for platform-specific factors that may affect both TVL and blockchain-related disclosure. First, I obtain from DeFi Pulse the dates which the platform's smart contracts were audited and construct a dummy variable that takes the value of 1 after the platform has been audited (*Audit*)[12]. This accounts for the smart contract code robustness as platforms that have been audited are deemed safer than platforms that have not been audited. I also control for the platform's age (*Age*) by taking its first tweet as its starting age. Controlling for age is important as platforms with a longer history may have higher TVL than newer platforms. In addition, I include market-related controls from Coinmarketcap, which is the most common data source used for cryptocurrency research (Lyandres et al. 2022). The market-related controls pertain to token

[10] Prior accounting literature examining earnings tweets by S&P1500 firms typically use a keyword search to identify earnings tweets. However, a similar approach is less feasible in the DeFi setting. For example, if we were to define blockchain-related tweets using the keyword "blockchain", we can see from Appendix B1 that it will likely not capture any blockchain-related topics but instead capture the non-blockchain related topic of "Strategic Partnerships". Sample tweets in Appendix B2 illustrates that using a keyword approach is likely to yield inaccurate results when identifying blockchain-related tweets. Rather, Twitter-LDA uses the underlying latent meaning of text to more accurately group similar tweets together and avoids the measurement error of a keyword list approach.
[11] Appendix B2 presents sample tweets post by DeFi platforms, organized by LDA topic.
[12] While DeFi Pulse requires all platforms to be audited, the website includes backfilled TVL data during which the platform may not have been audited.

characteristics as most DeFi platforms issue a platform token that gives token holders the right to vote on governance issues.[13] Platform tokens thus derive their value from the underlying economic activities on the platform, rather than from discounting cash flows as in standard valuation models (Cong et al. 2021). For these platform tokens, I calculate the token price volatility (*Token Volatility*) to control for fluctuations in the platform's economic activities, market capitalization (*Market Cap*) to control for platform size, and token trading turnover (*Token Turnover*) to control for platform token liquidity. Since not all platforms issue tokens, I also include a dummy variable for whether that platform has a token listed on Coinmarketcap (*Token Dummy*) and assign a value of 0 for market-related controls if that observation has missing token-related data. I winsorize all continuous variables at the 1st and 99th percentile.[14]

3.3 Research Design

To test H1 of whether voluntary blockchain-related disclosure affects the platform's TVL, I estimate the following regression for each platform *i* and month *t*:

$$Ln\ (TVL)_{i,t} = \alpha + \beta_1 Ln\ (Blockchain\ Tweets)_{i,t-1} + \beta_2 Audit_{i,t-1} + \beta_3 ETH\ Momentum_{i,t-1} + \beta_4 Age_{i,t-1} + \beta_5 Token\ Dummy_{i,t-1} + \beta_6 Token\ Volatility_{i,t-1} + \beta_7\ Market\ Cap_{i,t-1} + \beta_8 Token\ Turnover\ _{i,t-1} + Platform\ FE + Year\ FE + \epsilon \qquad (1)$$

[13] For example, holders of the platform token issued by Aave can vote on items such management of the platform, service improvement proposals and allocation of funds. In addition, token holders can also lend their tokens to the platform to cover lenders in case of a capital shortage, and receive platform fees in return (https://www.finder.com/aave).
[14] Please refer to Appendix C for all variable definitions.

I use the lagged value of *Ln(Blockchain Tweets)* in month t-1 to mitigate reverse causality concerns. Accordingly, all control variables are measured in the lagged period. In addition, I include platform fixed effects to control for time-invariant platform characteristics and year fixed effects to account for time-specific events affecting DeFi platforms. Standard errors are clustered at the platform level for all my analysis. If disclosure of summarized blockchain information is not helpful for users, then we would expect β_1 to be insignificant. However, if blockchain-related disclosure is useful for DeFi users and increases platform credibility, then we would expect a significantly positive β_1.

To test H2 on whether blockchain-related disclosure increases TVL due to information processing costs, I conduct cross-sectional tests using equation (1). I first partition on the lines of smart contract code (*Lines of Contract Code*) that the platform uses to proxy for direct costs that users incur for understanding the code and also to proxy for indirect processing costs as longer code is correlated with more complex platform structure. The second partition uses the lines of code that DeFi Pulse uses to extract the TVL from each smart contract (*Lines of TVL Code*). More lines of code correspond with the difficulty of extracting smart contract information from the platform. These cross-sectional tests aim to show that disclosure matters more when processing costs are higher.

4. Results

4.1 Descriptive Statistics

Table 1 Panel C illustrates the growth in DeFi platforms across my sample period and Figures 1 to 3 provide a graphical representation of the same growth trend. In 2018, there were only 6 DeFi platforms with a TVL of US$3 billion and 49,000 users. This increased sharply between 2020 and 2021, where the number of platforms grew from 69 in 2020 to 106 in 2021. Accordingly, TVL grew from US$16 billion to US$79 billion and the number of users from 1.7 million to 3.6 million over the same period. While the number of users seems to be decreasing starting in 2021 (see Figure 3), TVL still continued to increase. This suggests growing concentration of DeFi activity among groups of users, which echoes the findings of recent empirical papers (Cong et al. 2022). Of the different DeFi platform categories, lending platforms consistently have the largest TVL and number of users although the number of lending platforms does not dominate the sample. This suggests that most DeFi activity comes from borrowing and lending transactions. Another category that has grown exponentially is assets. These platforms mostly maximize returns for their users through yield maximizing strategies, such as shifting cryptocurrency assets between lending platforms to earn the highest yield. Such asset management services are similar to how fund managers manage assets in traditional finance.

The summary statistics for the main variables are reported in Table 2 Panel A. DeFi platforms have an average TVL of US$671 million across my sample period but are highly skewed with a standard deviation of US$2.1 billion. Thus, I use a log transformation of TVL to estimate my regression models. Of the 3,801 monthly unique users, 64% are new users, 25% are returning users, and only 11% are resurrected users. This suggests that DeFi users mainly continue using platform’s services or switch to another platform, but do not often switch back to platforms that they have previously used. DeFi platforms are also fairly active on Twitter, posting an average of 24 blockchain-related tweets per month, with an average of 35 words per tweet. Figure 4 shows that the tweeting behavior is largely similar across different DeFi categories and over the years. The average age of DeFi platforms is just above two years old (2.03), reflecting the relative immaturity of DeFi markets.

Next, I report the pairwise correlation for my main variables and highlight in bold the correlations that are significant at the 1% level. TVL is positively correlated with the lagged number of tweets and most lagged control variables. The number of blockchain-related tweets is positively correlated with contract code audit, token volatility, market capitalization and token turnover, suggesting that large platforms and platforms with greater trading and fluctuations in economic activities tweet more. In addition, platforms tweet more during high cryptocurrency sentiment (i.e., *ETH Momentum*), suggesting that platforms respond to DeFi users' demand for information by disclosing more frequently on Twitter.

To understand what drives DeFi platforms' tweeting behavior, I regress the number of blockchain-related tweets on several variables (results presented in Appendix D). Focusing on column 2 which includes all variables, I find that lagged ETH Momentum, token volatility and token turnover are positively associated with the number of blockchain-related tweets while age and market capitalization are negatively associated with the number of tweets. This suggests that platforms increase their voluntary blockchain-related disclosure on Twitter in response to users' demand for information (when cryptocurrency market sentiment and trading volume are high) and that younger and smaller firms tweet more. Thus, platforms respond to users' demand for information by posting more tweets and Twitter is an important disclosure channel in the DeFi market.

4.2 H1: Impact of Voluntary Blockchain-Related Disclosure on TVL

I test H1 by estimating the impact of the number of blockchain-related tweets on TVL – if disclosure helps users process blockchain information and assess platform credibility, then we should observe a significant association between the two. Table 3 shows that the number of blockchain-related tweets in the previous period is positively associated with TVL both when I omit and include fixed effects (columns 1 and 2 respectively). Next, I include market-related controls of platform tokens in column 3 and find that the positive effect remains significant. The coefficient of 0.4261 suggests that an increase in one standard deviation of platforms' tweets is associated with a 39% increase in TVL (i.e., US$261 million).[15] This suggests that the effect that blockchain-related tweets have on TVL is economically significant. In addition, I use an alternative way of measuring Twitter disclosure by replacing the number of blockchain-related tweets with the number of words used in blockchain-related tweets and present this result in column 4. The number of words used in blockchain-related tweets remains positively associated with TVL.

These results support H1 and suggest that platform's blockchain-related disclosure on Twitter matters for DeFi users. Even though DeFi platforms are programmed to be transparent, it appears that users cannot fully process public information on their own and have to rely on platforms' disclosure. Such disclosure facilitates users' assessment of platform credibility and thus increases users' confidence in and financial commitment to these platforms. One possible question that may arise is why platforms do not simply disclose

[15] (1 + (28/24)) ^ 0.4261 = 1.390

more to increase their TVL. First, since the average age of DeFi platforms is only two years old, DeFi markets are immature and it is possible that platforms have not learned the positive association between disclosure and TVL. Second and more importantly, platforms incur high processing costs to extract and summarize blockchain information. Unlike traditional firms with mature financial reporting systems in place, DeFi platforms likely do not have ready information on hand. Instead, DeFi platforms must access the public blockchain and programmatically extract relevant information, summarize it into understandable formats and then disclose them on Twitter. Such processing costs may deter some platforms from extracting and disclosing information frequently.[16]

4.3 How Information Content of Tweets Matters

After establishing that blockchain-related tweets have an impact on TVL, I next explore the type of information content that matters for this relationship. There are two variables I use to quantify information content. First, tweets providing less repetitive and more varied content should be more informative for DeFi users. When analyzing firms' 10-K filings, Dyer et al. (2017) shows that repeated sentences across documents can be redundant and have less information content for readers. Similarly, I argue that platforms posting blockchain-related tweets with dissimilar textual content within the

[16] Besides platform's voluntary disclosure, the DeFi market may also have other information intermediaries that could help users process and summarize blockchain information. One such example is DeFi Safety, which is an independent rating agency that evaluates the platform's process quality and assigns an overall safety score to platforms. To ensure that my results are not driven by such information intermediaries, I repeat my main test by adding the DeFi Safety score as a control variable and find that my results still hold (untabulated). I do not include the safety score in all my tests because it is not time-varying and thus including it would be absorbed by platform fixed effects. Thus, for this additional test, I include category instead of platform fixed effects.

same month provide more information for users. I capture the extent of repetition and variations in textual content by employing a machine learning technique developed by Google in Cer et al. (2018), Universal Sentence Encoder (USE), and used in the accounting literature by Crowley et al. (2021). USE captures the extent of content difference between two sentences and Appendix E shows an example. *Content Difference* represents the normalized Euclidean distance between all tweets posted by the platform within the same month using the USE algorithm. I partition my sample based on *Content Difference* and columns 1 and 2 of Table 4 show that the relationship between tweets and TVL is significantly stronger when *Content Difference* is larger.

The second information content variable focuses on the number of Twitter-LDA topics that blockchain-related tweets cover (*No. of Topics*). Platforms post an average of 3 blockchain-related topics during the month and the intuition follows that the more blockchain-related topics that the number of blockchain-related tweets cover, the greater the information content. I count the number of topics used by platforms within the month and partition my sample using the median of *No. of Topics*. Columns 3 and 4 show that the relationship between blockchain-related tweets and TVL is significantly larger when platforms post about a larger number of topics. Taken together, these results strengthen the main inference that voluntary disclosure is informative for users and thus increases platforms' TVL.

4.4 H2: Are Results Driven by Information Processing Costs?

H2 argues that blockchain-related disclosure increases TVL due to high information processing costs of public blockchain information. To test this, I conducted three cross-sectional tests that vary information processing costs for DeFi users. First, I directly obtain the smart contract code that DeFi platforms rely on to provide their services. For more technical DeFi users who can understand and process programming code, the more lines of code that the platform uses in their smart contract, the higher the direct processing costs incurred to understand and evaluate the smart contract. In addition, for general DeFi users, the lines of smart contract code capture the complexity of the platform's structure and thus proxy for users' indirect processing costs. For example, a platform offering a single financial service likely requires fewer lines of code to execute its smart contract as compared to another platform offering multiple financial services. Users thus incur higher processing costs to evaluate the platform with multiple financial services. If voluntary disclosure is informative for DeFi users due to high processing costs, we should expect to see a stronger relationship between blockchain-related tweets and TVL for platforms using more lines of contract code. Columns 1 and 2 of Table 5 show that the main results are driven by the sub-sample of platforms with a higher than median *Lines of Contract Code* and that the difference between sub-samples is significant. This supports my argument that platforms using many lines of code cause their users to face high processing costs and thus these users increase their reliance on platform disclosure.

Next, I use another indirect measure of information processing cost by exploiting DeFi Pulse's open-source code. To calculate each platform's TVL, DeFi Pulse extracts and sums token balances from platforms' smart contracts. The more complex the platform's services are, the more lines of code are needed to extract these token balances. For example, the lending platform, Aave, employs a multi-layer complex structure to provide financial services and users must process the multiple smart contracts embedded within the structure to obtain Aave's TVL. In contrast, another lending platform, Reflexer, employs a simple structure that makes it more straightforward to obtain its TVL. DeFi Pulse uses 457 lines of code to calculate Aave's TVL but only 20 lines of code to obtain Reflexer's TVL. I thus partition my sample based on the lines of code used by DeFi Pulse to calculate each platform's TVL (*Lines of TVL Code*), where a higher than median *Lines of TVL Code* suggests that users likely incur higher costs to parse through and evaluate the platform's complex structure. Similar to the results using the *Lines of Contract Code*, the relationship between tweets and TVL is driven by the sub-sample with greater *Lines of TVL Code* as shown in Table 5 columns 3 and 4, with the difference between sub-samples being significant. Both the results from using the platform's smart contract code and DeFi Pulse's TVL code support my argument that voluntary blockchain-related disclosure is helpful for users when the platform has high information processing costs. Taken together, the cross-sectional tests provide support for H2 that the association between blockchain-related tweets and TVL is due to information processing costs.

4.5 Alternative Measure of TVL: Number of DeFi Users

Although TVL captures DeFi users' financial commitment to the platform and is often used as a measure of user confidence, one limitation is that it can also be affected by the value of the underlying cryptocurrency assets even when real transaction activities do not change (see Appendix A). I thus use an alternative measure to proxy for user confidence in the platform. If blockchain-related tweets lower processing costs of blockchain information, then a greater number of users should be able to understand blockchain information about the platform and have greater confidence in transacting with the platform. To engage the platform's financial services, DeFi users have to use their cryptocurrency wallets with a unique wallet address (conceptually similar to a bank account number) to complete the transactions. I thus use the unique number of wallet addresses transacting with the platform's smart contract to proxy for the number of users. Doing so also gives rise to a nice feature where we can identify if a wallet address is interacting with the platform for the first time (i.e., new user), for a consecutive recurring time (i.e., returning user), or for a non-consecutive recurring time (i.e., resurrected user).[17]

Columns 1 and 2 of Table 6 show that the number of blockchain-related tweets in the previous period is positively associated with the number of unique platform users both when I omit and include fixed effects. Furthermore, columns 3 to 5 present the breakdown by user type and show that blockchain-related tweets increase both new and returning users. This suggests that blockchain information disclosed on Twitter is helpful not just for potential platform users, but also for existing users who may face high costs in continually processing blockchain information and thus rely on summarized blockchain-related disclosure instead. In contrast, blockchain-related disclosure does not have a significant impact on resurrected users, who are users that have previously interacted with the platform but did not in the previous month. This

[17] I do not use the number of users for my main tests due to the following two reasons: (1) TVL is empirically closer to my conceptual construct of interest as it represents users' financial commitment to the platform whereas number of users may not fully capture users' commitment and confidence in the platform. (2) Using the number of wallet addresses to proxy for the number of users also has limitations in that each user can own multiple wallets.

may be because such users return to platforms for reasons other than information processing costs.

Besides directly regressing the number of users on blockchain-related tweets, I also use this variable to check for extensive vs. intensive growth of TVL. My main results show that disclosure is positively associated with TVL and this could be driven either by a greater number of users transacting with the platform (i.e., extensive growth) or each user increasing their financial commitment to the platform (i.e., intensive growth). If disclosure indeed helps users process blockchain-related information, then ex-ante we should expect that disclosure increases TVL through enabling a greater number of users to understand and participate in the platform. Thus, I scale TVL by the number of unique users and regress this variable on the number of blockchain-related tweets. I find that tweets do not have a significant association with TVL scaled by users (untabulated). This suggests that tweets do not increase the intensive margin growth of TVL and is more likely driven by a greater number of users participating in the platform, which is consistent with my argument that tweets lower information processing costs for users. More specifically, blockchain-related tweets increase TVL through attracting more new users and retaining more existing users.[18]

[18] There are 3 possible scenarios that increases TVL: (1) greater number of new users (i.e., extensive), (2) retaining a greater number of existing users (i.e., extensive), (3) both new and existing users put a greater amount of investment in the platform (i.e., intensive). The test described here is meant to rule out the third explanation which is less consistent with the information processing cost explanation. The insignificant result documented here suggests that TVL growth is not driven by new and existing users investing more in the platform, but driven by attracting a greater number of new users and retaining a greater number of existing users.

5. Additional Tests and Robustness Checks

5.1 Spillover Effect from Centralized Cryptocurrency Exchanges Hacks

Unlike DeFi platforms, centralized cryptocurrency exchanges (CEXes) are opaque and their transaction activities cannot be found on the public blockchain. Thus, during adverse events such as cybersecurity hacks, CEXes may covertly grant preferential trading access to certain users while halting trades on other users' assets. In contrast, when DeFi platforms suffer from hacks, users have direct access to any subsequent blockchain transactions and can analyse how the hack occurred. For example, when a DeFi platform (Poly Network) was hacked and had $610 million stolen, DeFi users could scrutinize the contract code and figure out that the attack was possible because of the platform's mismanagement of access rights to its key smart contracts.[19] The opaque structure of CEXes, however, disallows public monitoring during a period when users are most concerned about the security of their cryptocurrency assets. Such opacity in times of cybersecurity hacks thus increases the salience of how important it is to have access to public blockchain information within the cryptocurrency community.

In this section, I explore whether and how CEXes hacks affect the importance of voluntary blockchain-related disclosure by DeFi platforms. Hacks make salient the risks arising from opacity as users are unable to monitor CEXes' transaction activity following a cybersecurity hack. I hypothesize that the increase in salience of risk affects DeFi platforms in two ways. First, existing DeFi users may switch to platforms that voluntarily disclose more blockchain

[19] https://www.reuters.com/technology/how-hackers-stole-613-million-crypto-tokens-poly-network-2021-08-12/

information as they are better able to understand how such platforms operate. Second, centralized cryptocurrency users may move their funds to DeFi platforms as DeFi platforms are built on the transparent public blockchain. When deciding between DeFi platforms, these centralized cryptocurrency users may choose platforms with greater voluntary disclosure as disclosure facilitates users' understanding and monitoring of the platform. Taken together, the increased importance of transparency after CEXes hacks may increase the effect that voluntary blockchain-related disclosure has on TVL.

I test this empirically by creating a dummy variable that takes a value of 1 if any CEX suffered a cybersecurity hack in the previous month (*CEX Hack*). This variable is then interacted with the main blockchain-related disclosure variable (*Ln(Blockchain Tweets)*). I include DeFi category (like industry) fixed effects because I want to examine whether DeFi users choose more transparent platforms within each category that provides similar services. Table 7 Column 2 shows a positive and significant coefficient on the interaction term between hacks and disclosure, suggesting that voluntary disclosure has a stronger effect on TVL when there is an increase in the importance that DeFi users place on understanding blockchain information. In addition, the negative coefficient on *CEX Hack* suggests that DeFi platforms who do not post any blockchain-related tweets during a hack suffer a decrease in TVL. Similarly, the interaction term between hacks and disclosure has a positive impact on the unique number of platform users, which suggests that transparent platforms can attract more users. This validates the conjecture that cryptocurrency users place a high emphasis on platform disclosure when there is an increase in salience of risks arising from opacity.

Furthermore, as cybersecurity hacks on CEXes are arguably exogenous to DeFi platforms but still within the relevant cryptocurrency ecosystem, the occurrence of hacks could provide an exogenous variation to the importance of voluntary blockchain disclosure on TVL. The results from this section could thus mitigate correlated omitted variable concerns and support my main result that voluntary blockchain-related disclosure has a positive impact on TVL.

5.2 Breakdown by Blockchain-Related Topics

To provide more insights into the type of blockchain-related tweets that users find informative, I decompose my main disclosure variable into the 7 blockchain-related topics that it constitutes (i.e., Items 1 to 7 of Appendix B1). I first separately regress TVL on each of the 7 topics and find that all 7 are positively associated with TVL (results untabulated). Next, I include all 7 topics in the same regression and find that the following 3 topics remain significant (results untabulated): Current Yield, Governance, and Platform Features.

While I do not make any ex-ante predictions about the type of blockchain information that may be helpful, I evaluate ex-post if these three topics are in line with my information processing cost explanation. First, current yield indeed has high information processing costs as it is tedious to calculate. A user would first have to scrutinize the platform's smart contract code to understand how yield is calculated and then query all the transactions that are relevant for calculations before arriving at a summary statistic of yield. Even though historical yield is also tedious, its insignificant result suggests that users are more interested in current than historical yield. Second, governance tweets typically pertain to soliciting users to vote on proposals to improve the platform's services. These proposals are typically conducted on-chain and without governance tweets, users would have to directly monitor the blockchain to know if there is an upcoming proposal. Governance tweets could also release the results of past voting proposals and without such tweets, users would have to query the blockchain to extract and parse the votes to arrive at the final voting result. Thus, it is possible that governance tweets also have high information processing costs. Third, platform features refer to operational changes made to the platform. As platforms are operated by smart contracts, such changes necessitate changes to the smart contract code. Without such tweets, users would have to constantly monitor the smart contract code and assess whether and how the code has changed.

On the contrary, users may have easier access to blockchain transactions and transaction costs through web interfaces such as Etherscan. These tweets typically re-produce information pertaining to specific blockchain transactions or the transaction costs that are directly shown on Etherscan. This suggests that the information within such tweets have lower information processing costs and are thus less informative for users. Overall, the results from the breakdown by blockchain-related topics suggest that users react more to blockchain-information that has higher information processing costs.

5.3 Robustness Checks

My inference that platforms' blockchain-related tweets increase TVL is potentially confounded by two endogeneity concerns: (1) reverse causality where TVL increases the number of blockchain-related tweets, and (2) unobservable omitted correlated variables. To deal with these issues, I follow the recommendations in Armstrong et al. (2022) to exploit the features of my setting and triangulate my results across different research designs and tests.

First, I acknowledge there is a possibility for reverse causality, where it is TVL that is driving the number of blockchain-related tweets instead of disclosure increasing TVL. Periods with high TVL may signal greater user interest in DeFi platforms, which in turn elicits higher demand for information. I first show that the number of tweets is not significantly affected by TVL in the previous period (Appendix D). Furthermore, this concern is more severe if I include advertising tweets in my independent variable as platforms may try to capitalize on high cryptocurrency sentiment to attract more users. However, my main blockchain-related tweets variable does not include advertising tweets and more likely contain actual blockchain-related information.[20] Nevertheless, I address such reverse causality concerns in two ways to strengthen my inferences.

I first measure users' information demand using Google searches of "Ethereum". As DeFi platform activity and TVL are affected by the value of ETH, it is likely that users have greater interest in DeFi platforms when there is growing interest in the Ethereum blockchain.[21] Thus, if it is mainly TVL driving tweets, then we should see a stronger association between blockchain-related tweets and TVL when user interest in Ethereum is high. I measure user interest from Google searches of "Ethereum" and partition my sample using the median of *Ethereum Google Trend*. Table 8 Panel A shows that my results are significant in both sub-samples and that the difference between the two is statistically insignificant. Furthermore, the magnitude of *Ln(Blockchain Tweets)*'s coefficient is similar between the two subsamples and to my main

[20] Please refer to Appendix B2 for examples of advertising tweets under topics of "community events" and "trading opportunities"

[21] Greater interest in the Ethereum blockchain suggests higher future ETH returns.

analysis in Table 3. This result suggests that my main inference is less subject to reverse causality concerns as the effect of blockchain-related tweets on TVL continue to hold when user interest and thus information demand is low.

Furthermore, I address reverse causality concerns by employing a lead-lag structure in all my analysis. As I regress the number of blockchain-related tweets in period t-1 on the TVL in period t, it is less likely that future TVL affects prior period disclosure. I also repeat my main analysis of regressing TVL on lagged blockchain-related tweets while controlling for lagged TVL and find that my results continue to hold (untabulated). In addition, my results are still meaningful even in the presence of such reverse causality concerns. If TVL is indeed driving the number of blockchain-related tweets, it suggests that platforms are reacting to DeFi users' demand for information and thus increase their disclosure on Twitter. Thus, both disclosure and Twitter as a disclosure channel remain important.

The second endogeneity concern is the presence of unobservable omitted correlated variables that affect both TVL and number of blockchain-related tweets such that the observed positive association is driven by the omitted variables instead of a true relationship between the two. These variables could pertain to macro-level factors such as a cryptocurrency bull market, or platform-specific factors such as the size of the founding team. Macro-level factors may boost user confidence in DeFi platforms and attract more deposits in a bull market regardless of how often the platform discloses information; at the same time, a higher number of blockchain-related tweets may simply reflect greater DeFi market activity during a bull market. In addition, platform-specific factors may be directly related to TVL as it proxies for their scale of operations.

A platform with only one founder is likely to operate on a smaller scale and thus have lower TVL. At the same time, this individual founder has less resources to post frequently on Twitter, resulting in a low number of tweets. Another confounding platform-specific factor is the actual information found on the blockchain. It is possible that users are reacting to the underlying blockchain information that happens concurrently with platform's disclosure on Twitter, instead of reacting to the tweet itself (this is a common limitation in the disclosure literature). Thus, the positive association that I document between TVL and the number of blockchain-related tweets may be explained by these correlated omitted factors, instead of a true relationship between the two. I mitigate this correlated omitted variable concern in several ways.

First, as the DeFi platforms in my sample operate on the Ethereum blockchain, I control for ETH price momentum and year fixed effects in all my regression models. This holds constant macro-level factors such as cryptocurrency market sentiment. Second, I include platform fixed effects (like including firm fixed effects in traditional capital markets research), which addresses the concern that my results are driven by time-invariant platform-specific factors. This implies that characteristics such as the DeFi platform founder's propensity to tweet are absorbed by the platform fixed effects and does not confound the impact that tweets have on TVL. Third, my results are similar at the daily level and when I regress the change in TVL on the change in the number of blockchain-related tweets (untabulated). These two research designs are less subject to omitted variables issues than an analysis at a longer time interval. Fourth, I use cross-sectional tests to show that my results vary with factors in line with my main argument. The relationship between blockchain-related tweets and TVL is stronger when tweets have greater information content and when users face higher information processing costs. Any correlated omitted variable would have to be correlated with these two factors.

Finally, I quantify the potential impact of omitted variables using the approach developed by Frank (2000) and recommended by Larcker and Rusticus (2010). This method derives the minimum correlations necessary for an omitted variable to render a significant result insignificant. I first derive the impact threshold for a confounding variable (ITCV), which is defined as the product of the (1) partial correlation between *Ln(Blockchain Tweets)* and the confounding variable that renders insignificant the coefficient of *Ln(Blockchain Tweets)*, and (2) the partial correlation between *Ln(TVL)* and the same confounding variable. The higher the value of ITCV, the less likely my results are driven by an omitted variable.

Table 8 Panel B shows that the ITCV for *Ln(Blockchain Tweets)* is 0.0436. Following Larcker and Rusticus (2010), I then compare the magnitude of ITCV to the impact (*Impact*) of the other control variables used in my main model. The *Impact* of a control variable is the product of its partial correlation with *Ln(Blockchain Tweets)* and its partial correlation with *Ln(TVL)*. I find that ITCV has a larger magnitude than all the other control variables' impact. This finding suggests that if omitted correlated variables exist, then the omitted confounding variable must have a larger impact on *Ln(TVL)* than any other control variable to render *Ln(Blockchain Tweets)* insignificant. This test thus suggests that the likelihood that an unobserved omitted variable is driving my main results is very small.

5.4 Possible Alternative Explanation of Advertising Effect

Another possible explanation for why blockchain-related tweets increase TVL is that tweets have an advertising effect. Prior literature has shown that firms strategically use advertising to influence stock prices (Madsen and Niessner 2019). These ads do not convey new information and are designed to attract investor attention. There are two possible types of advertising tweets that exist in my setting. The first is a pure advertising tweet: “Kyber Network is currently the most popular exchange to trade DAI on!”. These tweets do not have information content and are not included in my sample. The second type contains blockchain information but may also be construed as advertising: “Bancor has now an $SNX liquidity pool!”. This new feature is coded into the smart contract and users would be aware if they had been tracking changes to the smart contract code. As this second type of tweet processes and contains blockchain-related information, I include them in my sample. Cross-sectional results that tweets with greater information content have stronger associations with TVL also support my argument that these tweets help users process blockchain information.

In addition, I address this possible concern with two additional tests. First, this concern is particularly severe if the blockchain-related tweets do not correspond with contemporaneous blockchain information that users have to process.[22] An advantage of my research setting is that I can directly observe if there are operational changes made to the platform. Since platforms are fully

[22] See Appendix B2 tweet example under the topic of “Governance”. The tweet here describes how voting works when executed on a blockchain, but does not immediately correspond with new information that can be gathered from the blockchain. An analogy in traditional capital markets is if a firm discloses that shareholder proposals can pass with a majority number of votes. While such disclosure is informative, it is likely less informative than disclosing that a certain shareholder proposal has passed after receiving x% of votes.

executed by smart contracts, any changes to transaction terms will have to be made through changes in their code. I construct a variable, *Contract Code Change*, that takes a value of 1 if the platform changed its smart contract code during the month. Table 9 presents the results using only the sample when the platform has changed its code and show that my results continue to hold.[23] This mitigates the concern that the variable (*Ln(Blockchain Tweets)*) simply attracts user attention without conveying actual information. Second, I control for the number of non-blockchain related tweets to hold constant the effect of other tweets and find that my main results continue to hold (untabulated).

5.5 Potential Signaling Effect

The premise of my main hypothesis about why disclosure matters is that the high information processing cost of blockchain information creates information asymmetry as users are unable to directly process blockchain transactions. Consequently, voluntary blockchain-related disclosure could mitigate the asymmetry problem by either helping users to process blockchain information or to provide a signal that the platform is credible. The results presented in this paper support the information role of disclosure by showing that disclosure has a stronger effect when the tweets have greater information content. In addition, it is also possible that disclosure increases TVL by playing a signaling role. As DeFi platforms disclose information that is summarized from the blockchain, their disclosure is verifiable by anyone who can download and process blockchain information. Such verifiability of blockchain-related disclosure thus gives credibility to voluntary disclosure as a signal. The possibility of disclosure playing a signaling role is thus also consistent with my

[23] The sub-sample where platforms do not change their contract code shows an insignificant impact of tweets on TVL (untabulated).

main hypothesis that voluntary disclosure has an important role to play even in a transparent DeFi market. Such a possibility, however, has potential regulatory implications. Since platforms use voluntary disclosure as a signal of their credibility, mandating all platforms to disclose may de-value disclosure as a useful signal for users.

6. Conclusion

DeFi platforms provide a unique research setting as they are built on complete transparency – full transaction details and the smart contract code that governs these transactions are publicly observable from the blockchain in real-time. Despite such unprecedented transparency, regulators often lament the lack of disclosure regulation for these platforms. SEC Commissioner Crenshaw asserts that the absence of regulation disadvantages small DeFi users and hints at the possibility of disclosure regulation for DeFi platforms. This apparent puzzle in requiring transparent platforms to disclose public blockchain information can be explained by information processing costs. As blockchain data is granular and requires advanced programming skills for proper evaluation, DeFi users may ignore the disaggregated public information and rely on summarized platform disclosure. Thus, I find that platform's voluntary disclosure of blockchain information on Twitter is positively associated with the platform's TVL and unique number of users. The relationship between blockchain-related tweets and TVL is stronger when the tweets provide greater information content and when users face higher information processing costs. My paper suggests that voluntary disclosure helps users process information when the available public information is too disaggregated and costly to process.

In addition, I find that disclosure has a stronger effect on TVL when the cryptocurrency community is primed to focus on information transparency due to cybersecurity hacks. Platforms that disclose more blockchain-related information attract both a greater number of users and higher TVL. Overall, my results suggest that DeFi platforms can help users understand granular blockchain information by summarizing and disclosing them on Twitter. These findings on the current information environment of DeFi markets could be informative for regulators who are focusing their attention on how to design effective disclosure regulation.

References

Albrecht, A. J., and J. E. Gaffney. 1983. Software Function, Source Lines of Code, and Development Effort Prediction: A Software Science Validation. *IEEE Transactions on Software Engineering* SE-9 (6).

Armstrong, C., J. Kepler, D. Samuels, and D. Taylor. 2022. Causality Redux: The Evolution of Empirical Methods in Accounting Research and the Growth of Quasi-Experiments. *Journal of Accounting and Economics*.

Beck, M. J., M. Glendening, and C. E. Hogan. 2022. Financial Statement Disaggregation and Auditor Effort. *Auditing: A Journal of Practice & Theory* 41 (2).

Bertomeu, J., X. Martin, and I. Sall. 2022. Measuring Defi Risk. *SSRN E-Journal*.

Blankespoor, E., E. DeHaan, and I. Marinovic. 2020. Disclosure Processing Costs, Investors' Information Choice, and Equity Market Outcomes: A Review. *Journal of Accounting and Economics* 70 (2–3).

Blankespoor, E., G. S. Miller, and H. D. White. 2014. The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™. *Accounting Review* 89 (1).

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).

Buterin, V. 2015. On Public and Private Blockchains. *Ethereum Blog Crypto Renaissance Salon*.

Cer, D., Y. Yang, S. yi Kong, N. Hua, N. Limtiaco, R. St. John, N. Constant, et al. 2018. Universal Sentence Encoder for English. In *EMNLP 2018 - Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Proceedings*.

Chainalysis. 2022. *The 2022 Crypto Crime Report*.

Chen, S., B. Miao, and T. Shevlin. 2015. A New Measure of Disclosure Quality: The Level of Disaggregation of Accounting Data in Annual Reports. *Journal of Accounting Research* 53 (5).

Chen, Y., and C. Bellavitis. 2020. Blockchain disruption and Decentralized Finance: The Rise of Decentralized Business Models. *Journal of Business Venturing Insights* 13.

Christensen, H. B., E. Floyd, L. Y. Liu, and M. Maffett. 2017. The Real Effects of Mandated Information on Social Responsibility in Financial Reports: Evidence from Mine-Safety Records. *Journal of Accounting and Economics* 64 (2–3).

Cointelegraph. 2022. Does the Future of DeFi Still Belong to the Ethereum Blockchain? https://cointelegraph.com/news/does-the-future-of-defi-still-belong-to-the-ethereum-blockchain.

Cong, L. W., Y. Li, and N. Wang. 2021. Tokenomics: Dynamic Adoption and Valuation. *Review of Financial Studies* 34 (3).

Cong, L. W., K. Tang, Y. Wang, and X. Zhao. 2022. Inclusion and Democratization Through Web3 and DeFi? Initial Evidence from the Ethereum Ecosystem. *SSRN E-Journal*.

Crenshaw, C. 2021. DeFi Risks, Regulations, and Opportunities. *The International Journal of Blockchain Law* 1.

Crowley, R., W. Huang, and H. Lu. 2022. Discretionary Dissemination on Twitter. *SSRN E-Journal*.

Crowley, R., H. Wenli, and H. Lu. 2021. Executive Tweets. *SSRN E-Journal*.

DeFiPulse. 2021. How to Join the DeFi Pulse Leaderboard. https://www.defipulse.com/blog/how-to-join-the-defi-pulse-leaderboard.

———. 2022. DeFi Pulse - The Decentralized Finance Leaderboard. https://www.defipulse.com/.

Deshmukh, S., A. Geest, D. Gogel, D. Resas, and C. Sillaber. 2021. Decentralized Finance (DeFi) Policy-Maker Toolkit - White Paper. *World Economic Forum* (June).

Dyer, T., M. Lang, and L. Stice-Lawrence. 2017. The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64 (2–3).

Financial Stability Board. 2023. The Financial Stability Risks of Decentralised Finance.

Frank, K. 2000. Impact of a Confounding Variable on a Regression Coefficient. *Sociological Methods & Research* 29 (2).

George, B. 2022. Why TVL Matters in DeFi: Total Value Locked Explained. *CoinDesk*. https://www.coindesk.com/learn/why-tvl-matters-in-defi-total-value-locked-explained/.

Gogel, D. 2021. DeFi Beyond the Hype: The Emerging World of Decentralized Finance. *Wharton Blockchain and Digital Asset Project* (May).

Hu, D., S. Markovich, and V. Zhang. 2022. Information Processing in a Transparent Market: Evidence from a DeFi Protocol. In *ABFER*.

Koh, K., Y. H. Tong, and Z. Zhu. 2022. The Effects of Financial Statement Disaggregation on Audit Pricing. *International Journal of Auditing* 26 (2).

Larcker, D. F., and T. O. Rusticus. 2010. On the Use of Instrumental Variables in Accounting Research. *Journal of Accounting and Economics* 49 (3).

Lu, J. 2022. Limited Attention: Implications for Financial Reporting. *Journal of Accounting Research*.

Lyandres, E., B. Palazzo, and D. Rabetti. 2022. Initial Coin Offering (ICO) Success and Post-ICO Performance. *Management Science*.

Madsen, J., and M. Niessner. 2019. Is Investor Attention for Sale? The Role of Advertising in Financial Markets. *Journal of Accounting Research* 57 (3).

Makarov, I., and A. Schoar. 2022. Cryptocurrencies and Decentralized Finance (DeFi). *SSRN E-Journal*.

Schär, F. 2021. Decentralized Finance: on Blockchain-and Smart Contract-based Financial Markets. *Federal Reserve Bank of St. Louis Review* 103 (2).

SEC. 2023. Kraken to Discontinue Unregistered Offer and Sale of Crypto Asset Staking-As-A-Service Program and Pay $30 Million to Settle SEC Charges.

Shukla, S., O. Kharif, and J. Ossinger. 2022. Billions of Dollars Flee to Crypto's Decentralized Roots After FTX Collapse. *Bloomberg News*. https://www.bloomberg.com/news/articles/2022-11-18/billions-of-dollars-flee-ftx-woe-to-crypto-s-decentralized-roots.

Zetzsche, D., D. Arner, and R. Buckley. 2020. Decentralized Finance. *Journal of Financial Regulation* 6 (2): 172–203.

APPENDIX A

Detailed Explanation of Total Value Locked (TVL)

Total Value Locked (TVL) is defined as the total value of cryptocurrencies locked in a DeFi platform's smart contracts and is conceptually like liquidity in the traditional finance context. TVL is calculated by summing the token balances held by each of the platform's smart contract and then multiplying these balances by their price in USD. As the token balance for each type of DeFi platform represents different user activities, I separately discuss how TVL is calculated for each of the DeFi categories.

For lending platforms such as Compound, token balances in its smart contracts increase when lenders deposit cryptocurrencies to earn interest and when borrowers deposit cryptocurrencies to be used as loan collateral. For Compound, when borrowers take out part of the lenders' cryptocurrencies as loans, the token balance will be reduced. However, the net balance of TVL increases as the value of the collateral that borrowers lock in is typically higher than the value of the loan that they take out. TVL is thus calculated as the sum of all cryptocurrencies deposited minus the sum of all cryptocurrencies taken out as loans and then multiplied by their respective prices in USD. Conceptually, TVL in lending platforms is like total deposits in traditional banks. High TVL (bank deposits) suggests that the DeFi platform (bank) has high liquidity and can continue providing lending and borrowing services.

For decentralized exchanges (DEXes) such as Uniswap, token balances increase when liquidity providers deposit their cryptocurrencies to facilitate exchange activities on the platform and when users deposit their cryptocurrencies to swap for another denomination. Token balances thus decrease when users successfully swap their cryptocurrencies and they withdraw their desired denomination. Conceptually, TVL for DEXes is like the liquidity that traditional exchanges hold to facilitate currency swaps. A high TVL suggests that the DeFi platform has sufficient funds so that trades can be processed efficiently.

For assets management platforms such as Yearn Finance, token balances increase when users deposit their cryptocurrencies to optimize returns and decrease when users withdraw their funds. Yearn Finance's smart contract then uses these funds to strategically earn high interest by moving it across different lending services such as Aave and Compound. Thus, TVL in assets management platforms is like assets under management (AUM) in traditional funds, where a higher TVL signals a larger pool for investment strategies.

For derivatives platforms such as Synthetix, token balances increase when users deposit cryptocurrencies as collateral for synthetic assets like gold, Bitcoin, or USD. As the synthetic assets are collateralized by cryptocurrencies, the assets can be freely traded. Thus, TVL represents the collateral supporting the derivative market, which is a similar setup as in traditional derivatives trading.

For payment platforms such as Flexa, token balances increase when users deposit cryptocurrencies for the purpose of payments to merchants such as Starbucks. Consequently, token balances decrease when the payment is made. TVL here is like the balances held in digital wallets or traditional bank accounts and thus represents liquidity.

TVL is an important metric for DeFi platforms as it represents the financial commitment that users have made to a particular platform. Users must have high confidence in the continuing operations of the platform before they lock their cryptocurrencies into the platform's smart contract. With the deposited cryptocurrencies, platforms then have sufficient liquidity to keep their operations running. We can thus interpret TVL as similar to the assets that traditional firms use to provide goods and services. For example, a bank needs deposits to provide loans, and a manufacturing firm needs machines to produce goods. The more deposits a bank has and the more machines a manufacturing firm has, the more goods and services they can provide. A healthy level of TVL signals that the DeFi platform has sufficient liquidity and represents DeFi users' collective confidence in the platform as a going concern.

Despite its merits, TVL has also been criticized as being influenced by extraneous factors. Since most DeFi platforms operate on the Ethereum blockchain, TVL is heavily affected by fluctuations in the price of Ether (ETH), which are outside of the DeFi platform's control. For example, when ETH increases in value, the token balances on DeFi platforms also increase in value, which in turn results in a higher TVL when converted to USD. Thus, even without users depositing more cryptocurrency, TVL could increase simply because the value of the token balances in the smart contract has increased. I mitigate this concern in two ways. First, I control for price momentum of ETH in all my regression models to hold constant extraneous network factors that are likely outside the DeFi platform's control. Second, I include DeFi platform fixed effects to account for any time invariant factors that explain how the DeFi platform's token value is linked to extraneous factors.

APPENDIX B1

Key Words of 15 Twitter-LDA Topics

This appendix presents the results from Twitter-LDA results of DeFi platforms' tweets. I assign topic labels using the top key words and sample tweets from each topic. Appendix B2 further presents sample tweets for each Twitter-LDA topic.

S/N	**Topic Label**	**Top 5 Key Words**	**% of Sample**	**Blockchain Related?**
1	*Transaction Costs*	liquidity; trade; gas; fee; price	9%	Yes
2	*Current Yield*	pool; yield; earn; deposit; reward	7%	Yes
3	*Governance*	vote; proposal; governance; token; stake	6%	Yes
4	*Blockchain Transactions*	cover; eth; dai; defi; deposit	4%	Yes
5	*Historical Yield*	option; eth; price; call; trade	4%	Yes
6	*Platform Features*	bancor; token; liquidity; pool; stake	4%	Yes
7	*Financial Metrics*	wing; supply; apr; badger; apy	2%	Yes
		Total Blockchain-Related	**36%**	
8	*Platform Founders*	community; team; work; defi; build	13%	No
9	*Strategic Partnerships*	defi; crypto; index; blockchain; asset	10%	No
10	*Community Events*	join; community; ama; discord; defi	9%	No
11	*Customer Support*	idex; token; support; trade; issue	9%	No
12	*Trading Opportunities*	alpha; farm; reward; pool; stake	6%	No
13	*Platform Awareness*	liquidity; trade; reward; farm; earn	6%	No
14	*Contests*	nft; win; winner; trade; prize	6%	No
15	*Cryptocurrency Advertisements*	app; card; bitcoin; btc; crypto	5%	No
		Total Non-Blockchain Related	**64%**	

APPENDIX B2

Sample Tweets from Each Twitter-LDA Topic

S/N	Topic Label	Sample Tweets
1	*Transaction Costs*	“@0xScott @UniswapExchange Gas price shouldn’t be higher than 0.1 gwei”
2	*Current Yield*	“There are large yield dollar positions that are being liquidated. If you are a sponsor, please rush to add more collateral if possible. There are liquidation opportunities available through that same UI.”
3	*Governance*	“'How does meta governance work? The meta-gov voting period will end 24 hours before the underlying $COMP governance vote. $INDEX holders vote on Index Coop’s Snapshot page above. If a 5% quorum is reached, Index Coop will execute the vote in accordance with the majority”
4	*Blockchain Transactions*	“New #Ethereum #Loan Request on #ETHLend - Loan amount 0.1015731 $ETH - Profit 0.1015731 ETH - Backed by 80,000 $LEND Tokens.”
5	*Historical Yield*	“Weekly Performance of ETH Theta Vault: The vault sold the $2.2k Strike Call and ETH expired at ~$2.06k, so the options expired out of the money. The vault earned ~0.37% (~21% APY) of yield in ETH. Another great week for stacking ETH!”
6	*Platform Features*	“We’ve burned 40,000 STAKE tokens on Ethereum and added 2 new validator candidate nodes running @Nethermind to prep for public POSDAO!”
7	*Financial Metrics*	“Total supply on #LendfMe, the money market of @dForcenet, hits a new all-time high of $13M. Earn more #crypto by saving it.”
8	*Platform Founders*	“b'RT @avichal: good example of why vesting for founding teams matters. Also at using Litecoin as an example of a project that survived.”
9	*Strategic Partnerships*	“Sharpe Announces a Cooperation Partnership with ETHLend!! #Sharpe #ETHLend #blockchain”
10	*Community Events*	“Join #Alpha AMA today (Mar 17th) at 3pm UTC on Alpha Finance Lab discord.”
11	*Customer Support*	“@_MrHand_ Working well here. DM us if the issue persists.”
12	*Trading Opportunities*	“Alchemix lets you manifest your future yield into the present. We've been hearing from some community members about what they plan to finance with our self-repaying loans. What will you do with your $alUSD?”
13	*Platform Awareness*	“@duniacryptoid @FortuneTube @TraderTrek @anrubudda @Gomblo112 Thank you for the attention to ForTube, if you have any question you can contact ForTube telegram: https://t.co/UqM3RADyqA. ”
14	*Contests*	The end date of the #1inch #Discord #Meme #Contest moves to July 30! Got no potential winners so far. Cheaters shall not pass! Honest meme makers still have good chances to win.
15	*Cryptocurrency Advertisements*	“Every informed person needs to know about #Bitcoin because it might be one of the world’s most important developments by #LeonLuow. What’s your guess on this? We guess that he is right, Bitcoin and #Blockchaintechnology are one of the biggest developments.”

APPENDIX C

Variable	Description
Dependent Variables	
Ln(TVL)	The natural logarithm of total value locked (TVL) within the DeFi platform, denoted in USD, and averaged over the month DeFi Pulse calculates TVL by extracting the total balance of crypto assets held by the DeFi platform's smart contracts. TVL is a key metric[24] for gauging user confidence in that particular DeFi platform.
Ln(Unique Users)	The natural logarithm of the unique number of wallet addresses that interacted with the platform's smart contract within the month. Unique users is the sum of: (1) new users, (2) returning users, and (3) resurrected users, as defined below
Ln(New Users)	The natural logarithm of the unique number of new wallet addresses that interacted with the platform's smart contract for the first time.
Ln(Returning Users)	The natural logarithm of the unique number of returning wallet addresses that interacted with the platform's smart contract both last month and in the current month.
Ln(Resurrected Users)	The natural logarithm of the unique number of wallet addresses that interacted with the platform's smart contract more than a month ago, then stopped interacting last month, and continued interacting again in the current month.
Independent Variables of Interest	
Ln(Blockchain Tweets)	The natural logarithm of one plus the total number of blockchain-related tweets that the DeFi platform has posted in the month.
Ln(Total Words)	The natural logarithm of the total number of words used in *Ln(Blockchain Tweets)*
Control Variables	
Audit	Dummy that takes the value of 1 after a smart contract code audit has been conducted for the DeFi platform.
ETH Momentum	Ether (ETH) momentum using ETH returns compounded over the month.
Age	Platform age is defined as the number of years since the platform's first inception on Twitter.
Token Dummy	Dummy that takes a value of 1 if the DeFi platform has a platform token with price information available on Coinmarketcap and 0 otherwise.

[24] https://www.coindesk.com/learn/why-tvl-matters-in-defi-total-value-locked-explained/

Token Volatility	The annualized standard deviation of the platform token's returns over the month
Market Cap	The number of platform tokens in circulation multiplied by the closing price of the platform token as at the end of the month (scaled in millions)
Token Turnover	The total trading volume of the platform token during the month divided by the number of platform tokens in circulation (scaled in thousands)
Cross-sectional Variables	
Content Difference	The normalized Euclidean distance between all tweets posted by the platform within the same month using the Universal Sentence Encoder algorithm. The more textual content differs, the larger *Content Difference* is.
No. of Topics	The number of blockchain-related topics (as defined by Twitter-LDA) that the platform has tweeted about during the month.
Lines of Contract Code	The number of lines of code that the platform uses in its smart contract.
Lines of TVL Code	The number of lines of code that DeFi Pulse uses to extract the TVL from each DeFi platform.
Additional Test Variables	
CEX Hack	Dummy that takes a value of 1 if a centralized cryptocurrency exchange suffered from a cybersecurity hack.
NCSkew	The negative coefficient of skewness of daily returns in the subsequent 1- or 3-months
Avg Unique Users	The average of *Ln(Unique Users)* over the previous 3 months
Avg Blockchain Tweets	The average of *Ln(Blockchain Tweets)* over the previous 3 months
Ethereum Google Trend	The Google Trend Index of searches of "Ethereum" during the month
Contract Code Change	Dummy that takes a value of 1 if the platform changed its smart contract code during the month and 0 otherwise.

APPENDIX D

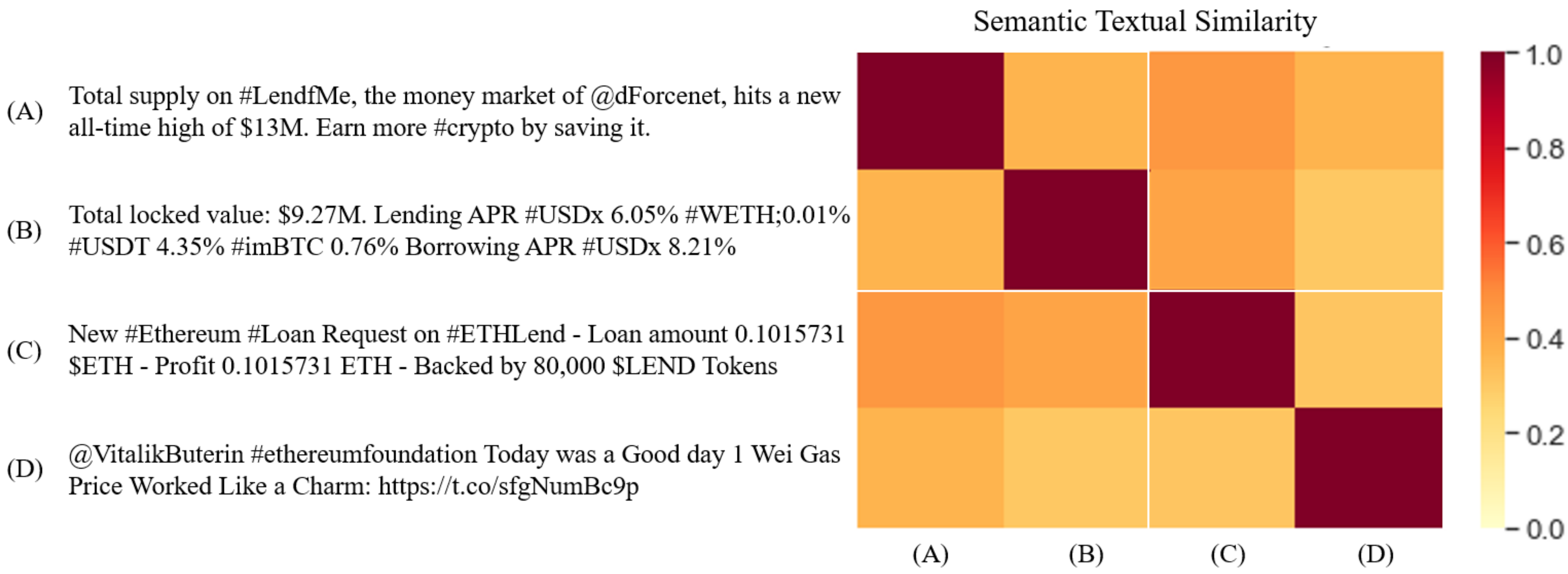
Determinants of Voluntary Disclosure on Twitter

This appendix reports the coefficient estimates from the determinants of platforms' voluntary disclosure on Twitter. Column (1) reports the estimates without market-related controls, while column (2) reports the estimates including all controls. The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(Blockchain Tweets)$_t$*	(2) *Ln(Blockchain Tweets)$_t$*
Ln(TVL)$_{t-1}$	0.0425	0.0400
	(1.4874)	(1.3388)
Audit $_{t-1}$	0.0840	0.0340
	(0.6505)	(0.2541)
ETH Momentum $_{t-1}$	0.1191**	0.1309**
	(2.1033)	(2.3070)
Age $_{t-1}$	-0.1345**	-0.1306**
	(-2.2945)	(-2.2005)
Token Dummy $_{t-1}$		-0.0326
		(-0.2875)
Token Volatility $_{t-1}$		0.0920***
		(3.5948)
Ln(Market Cap) $_{t-1}$		-0.0002*
		(-1.7774)
Ln(Token Turnover) $_{t-1}$		0.0065***
		(6.9319)
Platform & Year FE	Y	Y
Observations	2,150	2,150
Adjusted R-squared	0.6608	0.6664

APPENDIX E

Example of Universal Sentence Encoder Algorithm



<u>Explanation</u>:

In the example above, (A) and (B) are posted by a DeFi platform within the same month, while (C) and (D) are posted by another platform in another month. The USE algorithm correctly identifies that (A) and (B) are more similar in content than (C) and (D) are, as seen by the darker shades of orange. (A) and (B) disclose the financial performance of the platform, with (A) describing the total market supply and (B) disclosing various financial metrics. In contrast, (C) highlights a specific blockchain transaction, while (D) discusses about transaction costs. Thus, (C) and (D) provide greater information content than (A) and (B) do. Interestingly, the algorithm also identifies that (C) provides similar content to (A) in that they both suggest yield opportunities even though the two tweets contain no overlapping words besides “on”, “new”, and “by”.

FIGURE 1

Growth in Number of DeFi Platforms

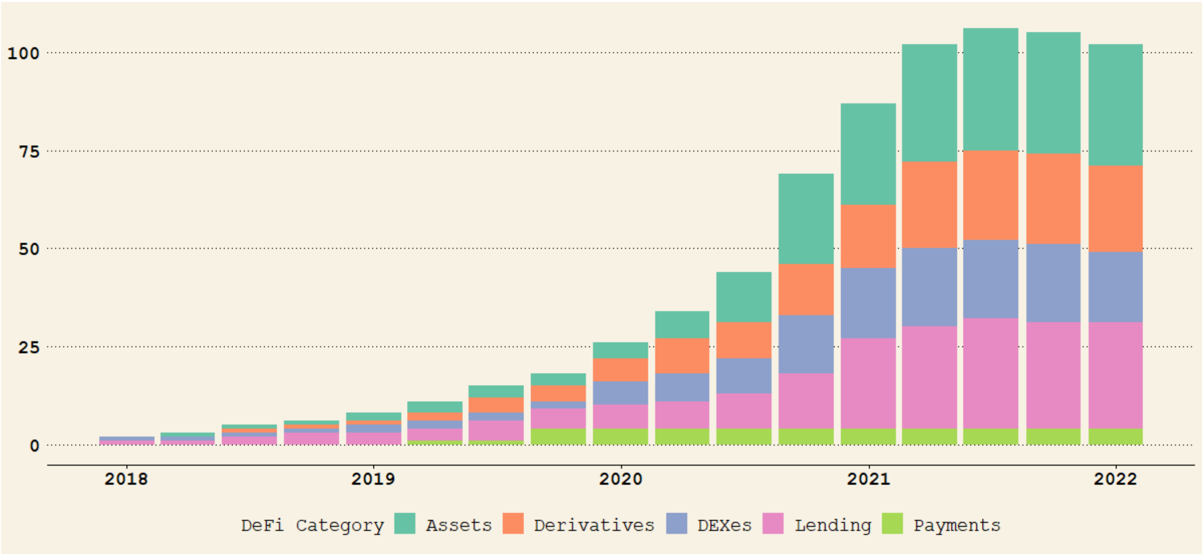


FIGURE 2

Growth in DeFi Platforms' Total Value Locked (USD)

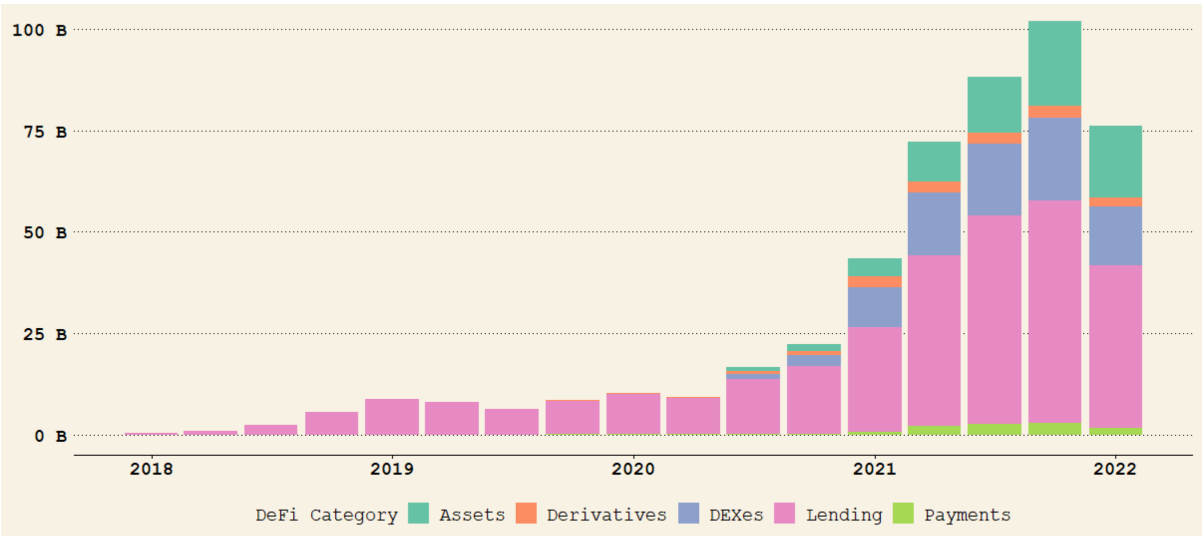


FIGURE 3

Growth in Unique Number of DeFi Users (in Millions)

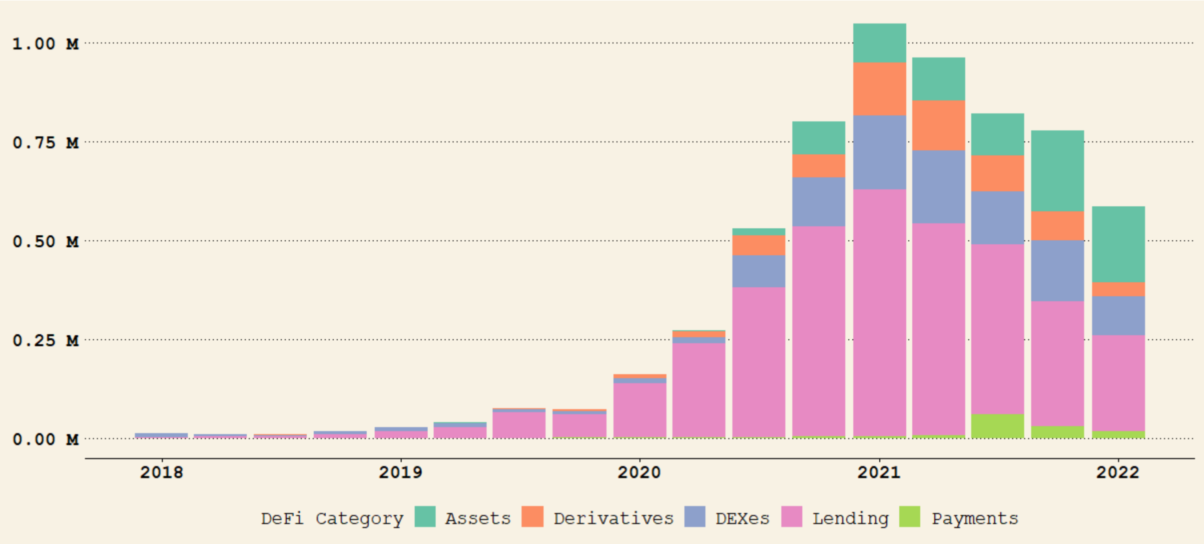


FIGURE 4

Time Trend of DeFi Platforms' Average Number of Blockchain-Related Tweets

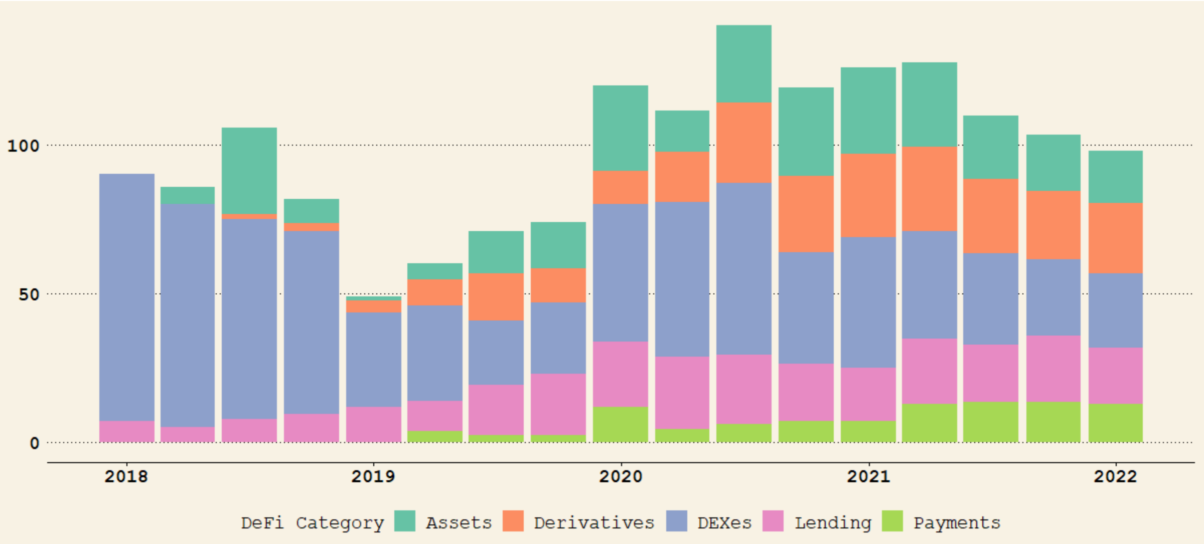


TABLE 1

Sample Selection and Distributions

This table presents the sample selection procedure (Panel A), annual distribution (Panel B), and distribution by DeFi category and year (Panel C). The final sample includes 2,150 platform-month observations between 2018 and 2022.

Panel A: Sample Selection

Year	No. of Platform-Month Observations	No. of Unique Platforms
Platform-months in DeFi Pulse Between Jan 2018 and Mar 2022	2,686	131
Less:		
Missing Twitter data	(339)	(10)
Missing control variables	(197)	(15)
Total	2,150	106

Panel B: Sample Distribution by Year

Year	No. of Platform-Month Observations	Percentage (%)	Cumulative Percentage (%)
2018	42	1.95	1.95
2019	144	6.70	8.65
2020	464	21.58	30.23
2021	1,193	55.49	85.72
Q1 2022	307	14.28	100
TOTAL	2,150	100	-

TABLE 1 – Continued

Panel C: DeFi Platform Statistics Distribution by Year and Category

Year	DeFi Platform Category	Total Unique No. of Platforms	Average TVL (USD billions)	Total Unique No. of Users (in Millions)
2018	Lending	3	2.982	0.019
	DEXes	1	0.025	0.027
	Assets	1	0.007	0.002
	Derivatives	1	0.002	0.002
	TOTAL	6	3.016	0.049
2019	Lending	5	7.744	0.166
	DEXes	2	0.009	0.034
	Assets	3	0.007	0.001
	Derivatives	4	0.061	0.008
	Payments	4	0.002	0.001
	TOTAL	18	7.823	0.210
2020	Lending	14	13.138	1.286
	DEXes	15	1.398	0.228
	Assets	23	1.098	0.102
	Derivatives	13	0.564	0.136
	Payments	4	0.093	0.008
	TOTAL	69	16.293	1.759
2021	Lending	28	45.050	1.906
	DEXes	20	16.337	0.658
	Assets	31	13.442	0.516
	Derivatives	23	2.760	0.425
	Payments	4	2.124	0.097
	TOTAL	106	79.714	3.602
Q1 2022	Lending	27	40.133	0.244
	DEXes	18	14.479	0.097
	Assets	31	17.749	0.192
	Derivatives	22	2.156	0.038
	Payments	4	1.670	0.015
	TOTAL	102	76.188	0.585

TABLE 2

Summary Statistics and Correlations

This table reports the summary statistics (Panel A) and pairwise correlations (Panel B) of the variables used in the analyses. Correlations significant at the 1% level are presented in bold font. Please refer to Appendix C for variable definitions.

Panel A: Summary Statistics

	N	Mean	St. Dev	p25	Median	p75
TVL (USD)$_t$	2,150	671,242,945	2,164,041,609	3,308,827	26,743,683	189,736,013
Unique Users $_t$	1,692	3,801.60	11,931.17	58.5000	320.00	1,984.00
New Users $_t$	1,692	2,340.99	7,353.15	22.0000	152.00	1,185.00
Returning Users $_t$	1,692	901.45	2,629.33	9.0000	85.0000	483.00
Resurrected Users $_t$	1,692	411.35	1,293.20	4.0000	29.0000	197.00
Blockchain Tweets $_{t-1}$	2,150	24.1014	28.8761	4.0000	13.0000	34.0000
Total Words $_{t-1}$	2,150	853.11	1,059.52	118.00	437.00	1,243.00
Audit $_{t-1}$	2,150	0.9316	0.2524	1.0000	1.0000	1.0000
ETH Momentum $_{t-1}$	2,150	0.1178	0.2772	-0.1258	0.0945	0.2965
Age $_{t-1}$	2,150	2.0302	1.8829	1.0000	1.0000	3.0000
Token Dummy $_{t-1}$	2,150	0.6772	0.4677	0.0000	1.0000	1.0000
Token Volatility $_{t-1}$	2,150	1.0122	1.0338	0.0000	0.8919	1.4508
Market Cap $_{t-1}$	2,150	177.67	468.01	0.0000	9.8705	111.65
Token Turnover $_{t-1}$	2,150	2.2076	16.6481	0.0000	0.0009	0.0241
Content Difference $_{t-1}$	2,150	0.7996	0.1690	0.8078	0.8746	0.9044
No. of Topics $_{t-1}$	2,150	2.9107	1.7191	2.0000	3.0000	4.0000
Lines of Contract Code $_{t-1}$	1,725	4,269.35	6,767.91	1,108.00	2,433.00	4,736.50
Lines of TVL Code $_{t-1}$	1,691	108.0426	97.23	42.00	77.00	148.00

TABLE 2 – Continued

Panel B: Pairwise Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) *Ln(TVL)* $_{t}$	1.0000										
(2) *Ln(Unique Users)* $_{t}$	**0.6917**	1.0000									
(3) *Ln(Blockchain Tweets)* $_{t-1}$	**0.2816**	**0.3795**	1.0000								
(4) *Ln(Total Words)* $_{t-1}$	**0.2483**	**0.3170**	**0.7396**	1.0000							
(5) *Audit* $_{t-1}$	**0.2161**	**0.1221**	**0.1803**	**0.1424**	1.0000						
(6) *ETH Momentum* $_{t-1}$	0.0432	**0.0728**	**0.0733**	0.0497	0.0481	1.0000					
(7) *Age* $_{t-1}$	**0.0955**	0.0466	**0.0534**	**0.0479**	**0.1833**	-0.0282	1.0000				
(8) *Token Dummy* $_{t-1}$	**0.2550**	**0.1606**	**0.1497**	**0.1453**	**0.4768**	0.0303	**0.2877**	1.0000			
(9) *Token Volatility* $_{t-1}$	**0.1317**	**0.1286**	**0.1702**	**0.1335**	**0.3032**	0.0421	**0.0750**	**0.6330**	1.0000		
(10) *Market Cap* $_{t-1}$	**0.4670**	**0.4243**	**0.1279**	**0.1261**	**0.1737**	0.0291	**0.2452**	**0.2905**	**0.1151**	1.0000	
(11) *Token Turnover* $_{t-1}$	**0.1017**	0.0574	**0.0532**	0.0428	**0.0615**	0.0189	**-0.0563**	**0.1010**	**0.0674**	**0.1799**	1.0000

TABLE 3

Impact of Blockchain-Related Tweets on TVL

This table reports coefficient estimates from OLS regressions of TVL on the number of blockchain-related tweets. Column (1) presents the results without fixed effects, while column (2) includes fixed effects. Column (3) additionally controls for market-related variables. Column (4) uses an alternative measure of disclosure, which is the total number of words used in blockchain-related tweets posted by DeFi platforms (*Ln(Total Words)*). The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(TVL)* $_t$	(2) *Ln(TVL)* $_t$	(3) *Ln(TVL)* $_t$	(4) *Ln(TVL)* $_t$
Ln(Blockchain Tweets) $_{t-1}$	0.6411***	0.4705***	0.4261***	
	(4.8652)	(4.7205)	(4.2635)	
Ln(Total Words) $_{t-1}$				0.1753***
				(3.5481)
Audit $_{t-1}$	2.4141***	1.0259***	0.8623**	0.8576**
	(5.4174)	(2.9937)	(2.4512)	(2.3697)
ETH Momentum $_{t-1}$	0.3838**	0.2455**	0.2788***	0.3011***
	(2.3194)	(2.3868)	(2.6345)	(2.7370)
Age $_{t-1}$	0.0997	0.4098***	0.3234**	0.3283**
	(0.6568)	(2.9327)	(2.3561)	(2.4519)
Token Dummy $_{t-1}$			0.6441**	0.6099*
			(2.0558)	(1.8964)
Token Volatility $_{t-1}$			0.0772	0.1217*
			(1.2042)	(1.9430)
Market Cap $_{t-1}$			0.0004	0.0004
			(1.5674)	(1.3666)
Token Turnover $_{t-1}$			0.0116***	0.0124***
			(4.4256)	(5.0414)
Platform & Year FE	N	Y	Y	Y
Observations	2,150	2,150	2,150	2,150
Adjusted R-squared	0.1264	0.7756	0.7848	0.7796

TABLE 4

Information Content of Blockchain-Related Tweets

This table presents results using the textual characteristics of tweets to measure information content. Cross-sectional tests are conducted by partitioning on the median of *Content Difference* (columns (1) and (2)) and *No. of Topics* (columns (3) and (4)). The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(TVL)* $_t$ High *Content Difference*	(2) *Ln(TVL)* $_t$ Low *Content Difference*	(3) *Ln(TVL)* $_t$ High *No. of Topics*	(4) *Ln(TVL)* $_t$ Low *No. of Topics*
Ln(Blockchain Tweets) $_{t-1}$	0.6333***	0.2740***	0.6358***	0.2645***
	(3.8820)	(2.6442)	(4.5498)	(3.3361)
Audit $_{t-1}$	0.2745	1.4128***	0.6325	1.0137**
	(0.8281)	(2.8560)	(1.6087)	(2.0160)
ETH Momentum $_{t-1}$	0.0565	0.4756***	0.1661	0.3582**
	(0.3555)	(2.9874)	(1.1752)	(2.2805)
Age $_{t-1}$	0.1651	0.3813**	0.3393***	0.2729
	(1.2073)	(2.2265)	(2.7816)	(1.3584)
Token Dummy $_{t-1}$	0.6703**	0.3968	0.5757	0.5754
	(2.1954)	(0.8133)	(1.5892)	(1.4626)
Token Volatility $_{t-1}$	0.0995	0.0320	0.1268	-0.0153
	(1.1465)	(0.3960)	(1.5466)	(-0.2240)
Market Cap $_{t-1}$	0.0007***	0.0005	0.0006**	0.0004
	(3.8249)	(1.4051)	(2.2803)	(1.3827)
Token Turnover $_{t-1}$	0.0169***	0.0127***	0.0094***	0.0210***
	(12.0406)	(3.4494)	(3.0687)	(4.2704)
Coef. Diff (p-value)	0.05*		0.01**	
Platform & Year FE	Y	Y	Y	Y
Observations	1,074	1,067	967	1,177
Adjusted R-squared	0.8017	0.7772	0.8139	0.7898

TABLE 5

Information Processing Costs

This table presents cross-sectional results when partitioning the main sample on information processing costs. Columns 1 to 4 use two variables: lines of smart contract code that the DeFi platform uses (columns (1) and (2)) and lines of code that DeFi Pulse uses to calculate TVL (columns (3) and (4)). The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1)	(2)	(3)	(4)
	Ln(TVL)$_{t}$	*Ln(TVL)*$_{t}$	*Ln(TVL)*$_{t}$	*Ln(TVL)*$_{t}$
	High *Lines of Contract Code*	Low *Lines of Contract Code*	High *Lines of TVL Code*	Low *Lines of TVL Code*
Ln(Blockchain Tweets)$_{t-1}$	0.5052***	0.0565	0.7330***	0.2128
	(4.3494)	(0.4955)	(5.5741)	(1.4640)
Audit$_{t-1}$	0.5868	0.4482	1.7449***	0.9646
	(1.1311)	(0.8951)	(4.1076)	(1.6136)
ETH Momentum$_{t-1}$	0.2374*	0.1404	0.1616	0.4008**
	(1.7178)	(1.0609)	(1.1129)	(2.3994)
Age$_{t-1}$	0.1160	0.6866***	0.2412*	0.4834**
	(0.6349)	(3.2578)	(1.8147)	(2.3694)
Token Dummy$_{t-1}$	0.5392	0.4836	1.4076***	0.0036
	(1.5021)	(1.2494)	(2.7677)	(0.0092)
Token Volatility$_{t-1}$	-0.0005	0.1636**	-0.1650*	0.1184
	(-0.0077)	(2.0152)	(-1.9971)	(1.2096)
Market Cap$_{t-1}$	0.0011***	0.0003	0.0008***	0.0008**
	(4.0330)	(1.1961)	(3.2028)	(2.5945)
Token Turnover$_{t-1}$	0.0106***	0.0250	0.0074***	0.0471
	(5.2912)	(0.9135)	(2.8757)	(0.6927)
Coef. Diff (p-value)	<0.01***		<0.01***	
Platform & Year FE	Y	Y	Y	Y
Observations	900	816	834	857
Adjusted R-squared	0.8170	0.8440	0.8380	0.7678

TABLE 6

Impact of Blockchain-Related Tweets on Number of Users

This table reports coefficient estimates from OLS regressions of the number of platform users on blockchain-related tweets. Column (1) presents the results without fixed effects, while column (2) includes fixed effects. Columns (3) to (5) show the breakdown by user type. The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(Unique Users)* $_t$	(2) *Ln(Unique Users)* $_t$	(3) *Ln(New Users)* $_t$	(4) *Ln(Returning Users)* $_t$	(5) *Ln(Resurrected Users)* $_t$
Ln(Blockchain Tweets) $_{t-1}$	0.6994***	0.2574**	0.2667**	0.3539***	0.0453
	(5.1739)	(2.2335)	(2.2707)	(3.0865)	(0.5066)
Audit $_{t-1}$	1.0339**	0.9153**	0.6073	1.4372***	1.7100***
	(2.0836)	(2.2423)	(1.3587)	(4.0782)	(5.1258)
ETH Momentum $_{t-1}$	0.4587***	0.3189***	0.4264***	0.0886	-0.0892
	(3.0692)	(3.1589)	(3.7389)	(0.9261)	(-0.7805)
Age $_{t-1}$	0.0348	-0.3541*	-0.4763**	-0.1423	0.1079
	(0.3627)	(-1.7216)	(-2.0790)	(-0.9879)	(0.7081)
Token Dummy $_{t-1}$		0.1030	-0.1726	0.6839**	1.4990***
		(0.3170)	(-0.4815)	(2.1758)	(4.7390)
Token Volatility $_{t-1}$		0.1614*	0.1793**	0.2126**	-0.2206***
		(1.9599)	(2.2795)	(2.3798)	(-3.4080)
Market Cap $_{t-1}$		0.0006**	0.0007***	0.0004**	0.0005**
		(2.4105)	(2.8806)	(2.0617)	(2.5464)
Token Turnover $_{t-1}$		0.0055	0.0006	0.0055	0.0107***
		(1.4756)	(0.0884)	(1.4522)	(5.9331)
Platform & Year FE	N	Y	Y	Y	Y
Observations	1,692	1,692	1,692	1,692	1,692
Adjusted R-squared	0.1593	0.7285	0.7030	0.7467	0.7710

TABLE 7

Spill Over Effect of CEXes Hacks

This table reports coefficient estimates from OLS regressions of TVL and platform users on blockchain-related tweets and *CEX Hack*. Column (1) presents the results using tweets and *CEX Hack* as standalone variables, while columns (2) and (3) include the interaction term between tweets and *CEX Hack*. The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(TVL)* $_t$	(2) *Ln(TVL)* $_t$	(2) *Ln(Unique Users)* $_t$
Ln(Blockchain Tweets) $_{t-1}$	0.6030***	0.5606***	0.5830***
	(5.4417)	(4.8666)	(4.8882)
CEX Hack $_{t-1}$	0.0795	-0.2973*	-0.2620*
	(1.3485)	(-1.7605)	(-1.6706)
CEX Hack $_{t-1}$ * *Ln(Blockchain Tweets)* $_{t-1}$		0.1493**	0.1235**
		(2.2387)	(2.0890)
Audit $_{t-1}$	1.3253***	1.3294***	0.3358
	(2.8331)	(2.8548)	(0.7978)
ETH Momentum $_{t-1}$	0.2484**	0.2487**	0.2820***
	(2.1932)	(2.1991)	(2.7660)
Age $_{t-1}$	-0.0356	-0.0341	-0.0243
	(-0.3801)	(-0.3634)	(-0.2663)
Token Dummy $_{t-1}$	0.8500**	0.8529**	0.2346
	(2.0180)	(2.0297)	(0.7033)
Token Volatility $_{t-1}$	-0.0936	-0.0917	0.0567
	(-1.1000)	(-1.0829)	(0.6247)
Market Cap $_{t-1}$	0.0024***	0.0024***	0.0019***
	(5.9540)	(5.9474)	(5.8560)
Token Turnover $_{t-1}$	0.0009	0.0010	-0.0041
	(0.1269)	(0.1435)	(-0.5038)
Category & Year FE	Y	Y	Y
Observations	2,150	2,150	1,580
Adjusted R-squared	0.4181	0.4186	0.3541

TABLE 8

Robustness Tests

This table first presents in Panel A the results from cross-sectional tests partitioning on the median of information demand measured from Google searches. *Ethereum Google Trend* is the total searches of "Ethereum" using the Google Trend Index. The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests. Panel B presents the results from assessing the impact of unobservable confounding variables using the method outlined in Frank (2000). For each independent variable, the impact threshold of confounding variable (*ITCV*) derives the minimum correlations necessary for an omitted variable to render the coefficient insignificant. *ITCV* is defined as the product of the (1) partial correlation between *Ln(Blockchain Tweets)* and the confounding variable that renders insignificant the coefficient of *Ln(Blockchain Tweets)*, and (2) the partial correlation between *Ln(TVL)* and the same confounding variable. *Impact* refers to the impact of the inclusion of each control variable on the coefficient of *Ln(Blockchain Tweets).*

Panel A: User Information Demand

	(1) *Ln(TVL)* $_t$ High *Ethereum Google Trend*	(2) *Ln(TVL)* $_t$ Low *Ethereum Google Trend*
Ln(Blockchain Tweets) $_{t-1}$	0.4218***	0.4363***
	(3.6186)	(4.2541)
Audit $_{t-1}$	0.9574**	0.8130**
	(2.0566)	(2.2537)
ETH Momentum $_{t-1}$	0.6875***	-0.1578
	(3.5091)	(-1.1761)
Age $_{t-1}$	-0.1793	0.4388***
	(-0.4071)	(2.9090)
Token Dummy $_{t-1}$	0.7818**	0.5985*
	(2.4826)	(1.7273)
Token Volatility $_{t-1}$	-0.0405	0.1309*
	(-0.4050)	(1.7702)
Market Cap $_{t-1}$	0.0005	0.0004
	(1.6367)	(1.4471)
Token Turnover $_{t-1}$	0.0099*	0.0120**
	(1.7499)	(2.0276)
Coef. Diff (p-value)	0.83	
Platform & Year FE	Y	Y
Observations	740	1,410
Adjusted R-squared	0.7593	0.7863

TABLE 8 – Continued

Panel B: ITCV Analysis

Dependent Variable =	*Ln(TVL)$_t$*	
	(1) *ITCV*	*(2)* *Impact*
Ln(Blockchain Tweets)$_{t-1}$	0.0436	
Audit $_{t-1}$		0.0066
ETH Momentum $_{t-1}$		0.0007
Age $_{t-1}$		0.0004
Token Dummy $_{t-1}$		0.0045
Token Volatility $_{t-1}$		0.0051
Market Cap $_{t-1}$		-0.0049
Token Turnover $_{t-1}$		0.0079

TABLE 9

Changes in Smart Contract Code

This table aims to examine tweets in relation to concurrent changes in blockchain information content. The main sample is restricted to observations where there is a concurrent change in the lines of smart contract code. Column (1) reports the coefficient estimates from OLS regressions of *Ln(TVL) on Ln(Blockchain Tweets)* without fixed effects, while columns (2) and (3) include fixed effects. The t-statistics are based on standard errors clustered at the platform level. Asterisks (***, **, and *) indicate statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-tailed tests.

	(1) *Ln(TVL)* $_t$ *Contract Code Change = 1*	(2) *Ln(TVL)* $_t$ *Contract Code Change = 1*	(3) *Ln(TVL)* $_t$ *Contract Code Change = 1*
Ln(Blockchain Tweets) $_{t-1}$	0.7861***	0.4222***	0.4056***
	(4.8008)	(3.7402)	(3.5930)
Audit $_{t-1}$	2.6953***	0.7421	0.6678
	(4.8785)	(1.4350)	(1.2906)
ETH Momentum $_{t-1}$	0.5177**	0.3525**	0.3642**
	(2.0762)	(2.1610)	(2.1686)
Age $_{t-1}$	-0.0763	0.4866**	0.4116**
	(-0.6891)	(2.2514)	(2.2436)
Token Dummy $_{t-1}$			0.5349
			(1.4568)
Token Volatility $_{t-1}$			0.0829
			(1.0490)
Market Cap $_{t-1}$			0.0008***
			(3.8224)
Token Turnover $_{t-1}$			0.0135***
			(5.0088)
Platform & Year FE	N	Y	Y
Observations	1,095	1,092	1,092
Adjusted R-squared	0.1497	0.7861	0.7981