MSc Finance – HEC Lausanne

Assignment

A Study on Decentralized Finance: What Are the Factors Influencing DeFi Lending and Borrowing Protocols?

Written and presented by:
Jérémie DE PINS DE CAUCALIERES
Charlie ORMOND
Nicolas HAMON
Maxime CARON
Class of 2024/2025

Learning Supervisor:

HURLIN Christophe

SUMMARY

INTRODUCTION

The rise of Decentralized Finance (DeFi) stands as a significant shift in the global finance ecosystem, offering accessible, transparent, and decentralized alternatives to traditional financial systems. Between 2016 and 2020 the market capitalization of DeFi increased dramatically from $600 million to $116 billion. *(Chainalasysis)*

Underdeveloped countries have been the main beneficiaries, with high adoption rates driven by limited access to traditional financial services. By allowing users to lend, borrow, and transact without intermediaries, DeFi contribute to financial inclusion and supports economic growth. As (*Biplab Kumar Guru and Inder Sekhar Yada, 2018)* highlight, the development of financial markets is essential for a country's economic progress, making DeFi a critical tool for underbanked regions.

In developed countries, institutional adoption is driving DeFi's growth. For example, the approval of Bitcoin and Ethereum ETFs by the SEC in 2024 reflects increasing efforts to integrate cryptocurrencies into mainstream finance. Political factors also play a role; the re-election of Donald Trump, a pro-crypto, highlights the push to reduce regulatory barriers and encourage adoption. *(Will Schmitt in New York, July 23, 2024)*

Figure 1 Global Crypto Adoption Index (Chainalysis, September 2022)

Among the various DeFi applications, lending and borrowing (L&B) protocols are central to its ecosystem. These protocols enable users to lock crypto-assets as collateral to borrow funds or earn interest, offering decentralized alternatives to conventional credit systems. Unlike traditional lending, DeFi L&B protocols cut credit checks, rely on algorithmically determined interest rates, and provide unparalleled transparency via blockchain technology. These features democratize access to credit while creating opportunities for financial inclusion.
Despite the importance of L&B protocols, there is limited understanding of what drives their performance. Metrics such as Total Value Locked (TVL), a key measure of the value of assets committed to these protocols, are often used as indicators of success. However, the role of factors like fees, token incentives, treasury size, and developer activity in influencing TVL stays under-explored. This study aims to address this gap by investigating the determinants of TVL across key L&B protocols, offering insights into what drives their growth and sustainability. (*jiahoua XU & Nikhil Vadgama, 11 January 2022)*

Figure 2 L&B index for 3 years (computed on Dune)

Figure 2 illustrates the progression of the L&B index, which currently stands at $6.7 billion, accounting for 5.77% of the DeFi market. To facilitate tracking the performance of individual sectors in DeFi, we have developed a comprehensive dashboard in SQL on *Dune Analytics*.

The structure of this paper is as follows:
Firstly, we provide an overview of the L&B sector, highlighting its risks and benefits compared to traditional finance.
Secondly, we detail the empirical methodology, including the panel regression framework and robustness checks.
Finally, we analyze the results, compare them with Monte Carlos simulation findings, and discuss their implications while suggesting directions for future research

The Lending & Borrowing sector:

Lending and borrowing platforms allow traders to deposit crypto assets as security and borrow other assets. Lenders earn returns on their security, while borrowers pay interest on their loans.

One of the defining features of DeFi lending platforms is their open-access nature. Users can lend or borrow assets without undergoing credit checks or providing extensive personal information, requiring only a digital wallet for participation. This streamlined process stands in deep contrast to traditional lending, where borrowers face rigorous credit evaluations, lengthy approval processes, and substantial documentation requirements.

Over-Collateralization and Interest Rates:

Most DeFi loans are over-collateralized due to the absence of credit checks. Borrowers must deposit assets, such as cryptocurrencies, that exceed the value of the loan they wish to obtain (e.g., depositing $150 worth of Ether to borrow $100 in stablecoins). This requirement acts as a risk management mechanism to protect lenders in the absence of traditional credit ratings.

Interest rates on DeFi platforms are dynamically determined by algorithms based on the supply and demand for assets within the protocol, unlike the fixed rates set by banks. *(Bartoletti, M., Chiang, J.H.Y., and Lluch-Lafuente A., 2020)*

Transparency and Privacy:

DeFi offers exceptional transparency, as all transactions, interest rates, and loan terms are recorded on the blockchain and accessible to the public. However, user privacy is preserved since DeFi platforms do not require personally identifiable information, allowing participants to remain pseudonymous. This contrasts sharply with traditional finance, where privacy is limited due to the need for identity verification and regulatory compliance. *(Maker DAO. (n.d.). Getting started: Maker Protocol 101)*

Risks and Challenges:

Despite its advantages, DeFi lending carries inherent risks. Smart contract vulnerabilities can be exploited by bad actors, potentially resulting in loss of funds. Furthermore, the volatility of crypto-assets can cause sudden changes in collateral value, leading to automatic liquidations. Regulatory uncertainty further complicates the DeFi landscape as governments are reticent with integrating these systems into existing financial frameworks. (*Betul Kaplan, Vahit Ferhan Benli, Elcin Aykac Alp, December 15, 2023)*

Table 1 centralized vs decentralized borrowing and lending methods

Feature	**Centralized (Traditional Fin)**	**Decentralized (Defi)**
Intermediary	Banks or financial institutions	Smart contracts on a blockchain
Approval Process	Credit checks and KYC	Permissionless, no KYC
Collateral	Often required, varies by loan	Over-collateralized
Interest Rates	Set by banks, can be rigid	Dynamic, algorithm-driven
Privacy	Requires extensive personal info	Pseudonymous, no personal data
Fees	Higher due to intermediaries	Lower, but include gas fees
Transparency	Limited, opaque to users	Fully transparent on-chain
Accessibility	Restricted by geography/regulations	**Global, accessible to anyone**

Empirical Context

Blockchain data is inherently public and open source, enabling comprehensive analysis of transaction flows. However, research often focuses on the price fluctuations of crypto assets rather than fully leveraging the available data. Tools like "Dune Analytics" provide access to blockchain data.

A study by Dominik Metelski and Janusz Sobieraj (2022) used causality analysis and panel regression to understand the valuations of DeFi protocols. The study highlighted TVL as a critical metric, representing the capital deposited in DeFi protocols.

Research by Corbet et al. (2021) employed Markov regime-switching vector regression analysis to explore the drivers of DeFi protocols, identifying significant impacts from assets like Chainlink (LINK) and Maker (MKR).

Other studies, such as Gilles Brice M'bakob's (2024) work on crypto-asset price changes, emphasized Metcalfe's Law and network effects as fundamental to blockchain performance. TVL remains a global metric for DeFi valuation, influenced by the values of native tokens and fiat currency fluctuations.

Overall, empirical evidence highlights the importance of TVL in DeFi markets, along with transaction volumes and network effects, reflecting user engagement and market dynamics.

Methodology

This study uses panel data regression applied to the L&B sector. As previously discussed, the TVL is a robust indicator of a DeFi protocol's valuation. TVL is calculated by summing the number of tokens locked within the protocol, which only accounts for the quantity utilized within the protocols, distinguishing it from market capitalization. Instead of using price as the dependent variable, we use TVL, which is much more correlated with the protocol's service and the real interest of users. For instance, the Ethereum network has a market capitalization of $360 billion, but only $55 billion worth of ETH is utilized for its services. Thus, 15% of ETH is actively used in the network, while the rest is retained by speculators for potential proftis. Since a higher TVL generally indicates a more valuable DeFi protocol, the question arises: which metrics significantly impacts TVL?

Data Sample

The data was extracted from Token Terminal and DeFi Llama, covering primarily the period from November 3, 2021, to November 19, 2023.

This study is based on four protocols: AAVE, MAKER DAO, COMPOUND, and CONVEX. Panel data regression was chosen for its ability to account for both time-varying characteristics and individual-specific factors, enabling a cross-sectional analysis that incorporates both individual and temporal effects. In panel data regression, three models can be employed: the pooled OLS model, the fixed effects model, and the random effects model. After conducting the necessary tests, the random effects model was found to be the most suitable for our data.

Descriptive statistics:

Secteur L&B	avg _TVL(million $)	avg _ cap(million $)	avg _FEES (annualized,million $)	avg _REVENUE(annualized,million $)	avg _Token price($)	avg _active user daily
AAVE	$ 1 100,00	$ 1 500,00	$ 389,00	$ 66,89	$ 100,00	10
COUPOUND	$ 1 800,00	$ 326,00	$ 22,33	$ 4,54	$ 40,00	433
CONVEX	$ 970,00	$ 250,00	$ 6,65	$ 2,74	$ 4,00	73
GOLD FINCH	$ 2,66	$ 177,70	$ 7,06	$ 0,76	$ 3,00	70
MAKER DAO	$ 6 750,00	$ 1 700,00	$ 230,00	$ 89,17	$ 2,70	113

Table 2 Descriptive statistics of Landing & Borrowing sector with 5 protocols

This table compares the four protocols in the L&B sector across various parameters:

avg_TVL : Indicates the amount of funds users have locked in each platform.
avg_cap : Represents the market value of each platform.
avg_FEES : Reflects the income platforms earn from fees.
avg_REVENUE : Represents the total income earned by the platform.
avg_Token price: Shows the average price of the platform's token.
avg_active user daily: Indicates the average number of daily active users on the platform.

Observation:

AAVE: Leads in revenue generation ($389 million) and has a high token price ($66.89), with 10 active users daily.

COMPOUND: Has moderate TVL ($1,800 million) and market cap ($326 million), but lower revenue ($22.33 million) and a token price of $40, with 433 active users daily, the highest among the listed platforms.

CONVEX: Has relatively lower metrics across all categories, with notable fees ($6.65 million) and a token price of $4.

GOLDFINCH: Shows the lowest TVL ($2.66 million), but a moderately high token price ($3.00). Revenue is the lowest ($0.76 million), with only 70 active users daily.

MAKERDAO: Dominates in TVL ($6,750 million) and market cap ($1,700 million), with significant revenue ($230 million), though the token price ($2.70) is relatively low. It has 113 active users daily.

General Trends:
Platforms with higher TVL or market cap do not necessarily lead in token price or active user count (MakerDAO vs. AAVE). User engagement (active daily users) does not always align with revenue or TVL (Compound has many users but lower revenue compared to AAVE).

Stationarity

To get an unbiased and consistent panel data regression we will have to check for stationarity in each variable:

"In its most intuitive sense, stationarity implies that the statistical properties of a process generating a time series do not change over time. This does not mean that the series itself remains unchanged, but rather that the manner in which it evolves remains constant over time. The algebraic equivalent of this concept could be a linear function, rather than a constant function; the value of a linear function varies with the increase in x*, but the manner in which this value changes remains constant — it has a constant slope, reflecting an unchanging rate of change" (Shay Palachy, 2019).*

In this study, we employ both temporal and cross-sectional models. Consequently, it is essential to test this assumption for each of the dependent and independent variables.

To do so, we must perform two tests: the Levin-Lin-Chu (LLC) test and the Augmented Dickey-Fuller (ADF) test. If the variables being tested are found to be non-stationary, we will address this issue by applying a differentiation process.

Differentiation in time series analysis is a method used to transform a non-stationary time series into a stationary one by focusing on the change (or difference) between successive values, rather than the values themselves. In this analysis, we have applied the first difference, which involved

$$\Delta y_t = y_t - y_{t-1}$$

subtracting the previous value from the current value for each independent variable, excluding the dependent variable (log TVL return), which is assumed to be stationary by nature.

Pooled (OLS) regression model, Fixed effect model (FE), Random effect model (RE):

The pooled regression model assumes the absence of unobserved individual heterogeneity among units, treating all observable data as fundamentally the same, which is not the case here as we can observe in the statistics descriptive section (page 7). The OLS model does not take in consideration either time effect or individual heterogeneity making this model obsolete for our model.

Yit = dependent variable

$$Y_{it} = \beta_0 + \sum_{k=1}^{K} \beta_k X_{k,it} + \alpha_i + \gamma_t + \epsilon_{it} \tag{4.1}$$

Xk,it = independent variables or predictor

αi = individual-specific random effect (which is the unobserved heterogeneity)

γt = time-specific effects, which capture factors that affect all individuals in the same way at a given time t

ϵit = error term

Due to all these reasons it is preferable to choose between a fixed effects model or a random effects model. We will run a Haussmann test to determine which model is more appropriate based on the nature of the data and assumptions about individual-specific effects.

Random effects are utilized when individual attributes (Xit) appear to have no correlation with the **unobserved individual effects** (αi). In other words, we assume that these individual effects impact the dependent variable, however they are **not correlated** with the independent variables "Cov(αi,Xit)=0". Also, estimates are based on both **within** and **between-individual variation**.

When running the Hausseman-test to check the significance between fixed effect or random effect. The test is not rejecting the null hypothesis indicating that the random effect is more significant for our model. And the assumption "Cov(αi,Xit)=0" still holds.

The random effect is a better choice not just because of the Haussmann test results but also because the properties fit our model and data:

- The model assumes that the individual-specific effects are randomly distributed across entities, which can be more realistic in many cases where the differences between entities are due to unobservable, random factors.
- The random effects allow for the incorporation of random variations in the intercepts.
- The random effects models make use of both within-entity and between-entity variation, they can potentially lead to smaller standard errors of the estimated coefficients.
- Compared to fixed effects models, random effects models may require fewer data points to achieve the same level of statistical power**.**

(Panel Data Analysis Fixed and Random Effects using Stata (v. 6.0) Oscar Torres)

Test for Bias in our selected model:

We are testing for three potential bias which are :

- Serial correlation or autocorrelation: meaning that the residuals of our model are correlated with each other. This is tested by proceeding to a "Durbin-Watson test."
- Heteroscedasticity: meaning that the variance of the residuals is not constant across the observations. We are using the "Breush-Pagan test."
- Cross-sectional dependence: when the residuals of the entities (independent variables) are correlated with each other. Two tests are used which are the "Pesaran CD test" and "Breush-Pagan LM test."

Our model shows significant Heteroscedasticity and Cross-sectional dependence, which needs to be corrected by some improvment. *(Greene W. 2018)*

Model improvement: "Cluster standard error":

The cluster standard error is used to adjust for **correlation within groups** or **clusters** in panel data or cross-sectional data where observations within each cluster may be correlated. It is mainly used to correct for autocorrelation, heteroscedasticity, and cross-sectional dependence, which are biases we have in our regression. It adjusts the variance-covariance matrix of the estimated coefficients to account for the intra-group correlation. .*(Greene W. 2018)*

Logarithmic Transformation:

A logarithmic transformation is applied to all variables since the variance of the dependent variable increases with its mean (heteroscedasticity). The logarithmic transformation stabilizes the variance, leading to more reliable and valid statistical inferences for each variable.

Handling Missing Values (N/A):

Missing values (N/A) are common and can render the dataset unusable. To address this issue, since the logarithm of zero is undefined, a small constant (e.g., 1) is added to the data before applying the logarithm. This ensures that zero values are transformed without causing errors, which is why "1" is added to all variables.

Results

Application: Lending and Borrowing

The best regression model is as follow:

$$\text{log\_return\_tvl}_{it} = \alpha + \beta_1 \cdot \text{diff\_log\_fees}_{it} + \beta_2 \cdot \text{diff\_log\_tokinc}_{it} + \beta_3 \cdot \text{diff\_log\_coredev}_{it} + \beta_4 \cdot \text{diff\_log\_treasury}_{it} + \mu_i + \lambda_t + \epsilon_{it}$$

(5.1)

Chosen Factors:

For each variable, a logarithmic transformation is applied to stabilize variance and mitigate the influence of outliers. The prefix "diff_" indicates that all independent variables have been transformed by taking their first differences.

Return of TVL_it: the return of the Total Value Locked for entities *i* at time *t* (TVL represents the total amount locked in a DeFi protocol in USD by users)

Fees_it: Fees for entities *i* at time *t* (the fees generated by the DeFi protocol)

Incentives_it: Token incentives for entities *i* at time *t* (the share of the DeFi protocol's governance tokens distributed to users)

Dev_it: Developers for entities *i* at time *t* (the number of developers working to ensure the proper functioning of the DeFi protocol)

Treasury_it: Treasury for entities *i* at time *t* (value in USD of the protocol's on-chain funds, including unallocated governance tokens).

μi : Entity-specific effect, indicator variables added to the regression corresponding to the number of factors

λt: Time-specific effect, dummy variable added to the regression corresponding to time

ϵit: Error term

Assumptions on how our independent variables will impact our model:

- **Fees:** Higher fees imply higher TVL.
- **Token incentives:** A greater availability of tokens for distribution tends to drive higher TVL.
- **Developers:** A larger number of developers contributes to a higher TVL.
- **Treasury:** A robust treasury can signify greater user confidence.

Tracing Procedure:

Initially, the collected metrics are plotted to visualize the entire dataset:

This chart visualizes the average log (TVL) (Figure 4) and average log (return TVL) (Figure 6) values for each protocol. The Figure (4) seems to be stationary but an additional test might be needed to identify its stationarity. However, the Figure (6) show less spread among observed data than graph 1.0 and seems more appropriate as dependent variables.

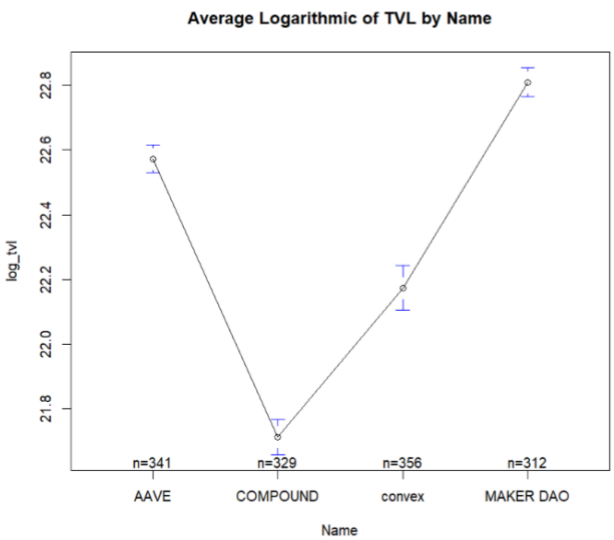

Figure 3

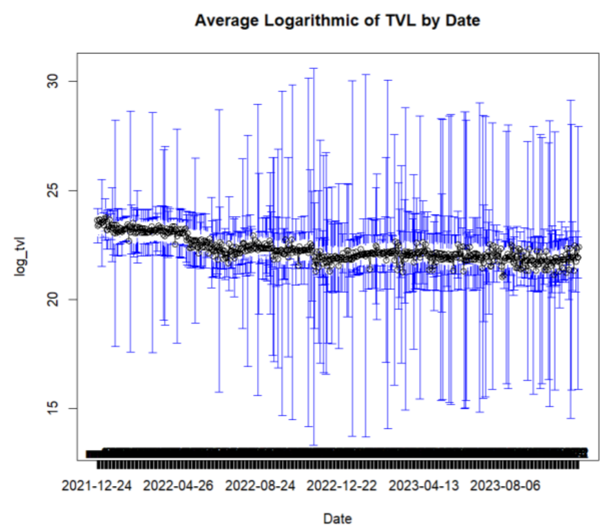

Figure 4

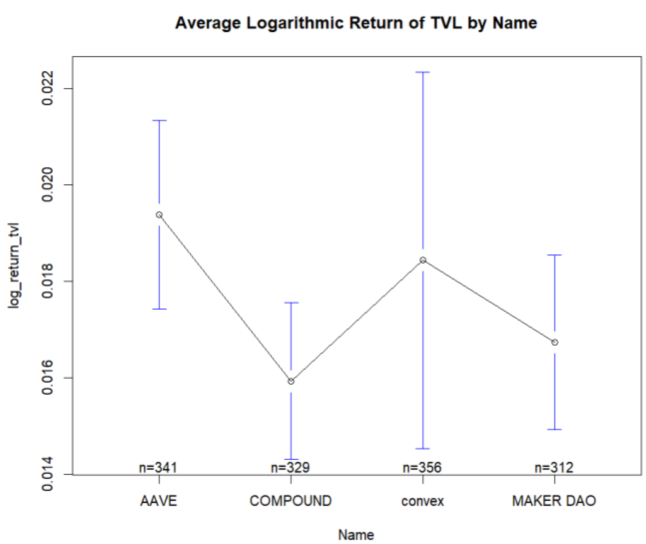

Figure 5

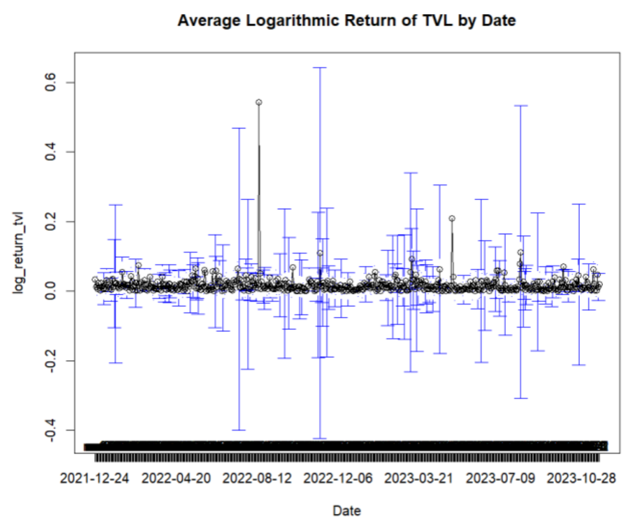

Figure 6

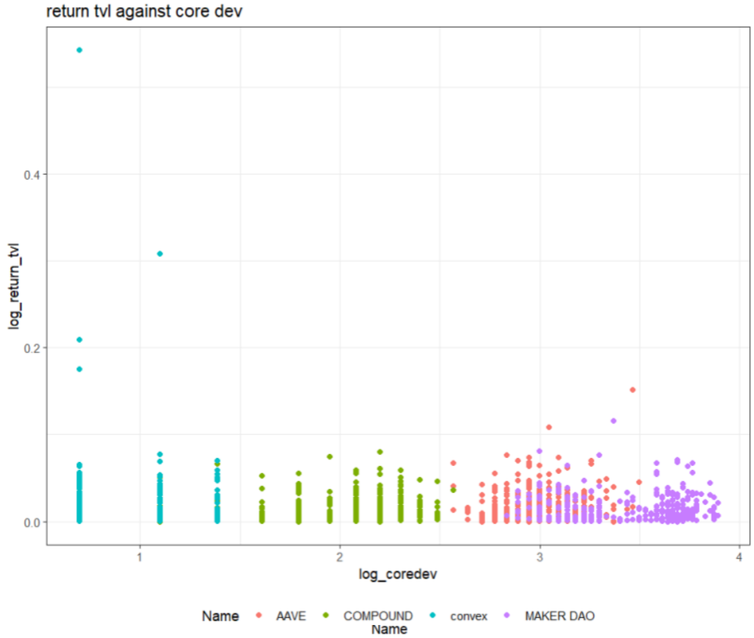

Figure 7 Dispersion of observed data between TVL and core dev (cross-sectional approach)

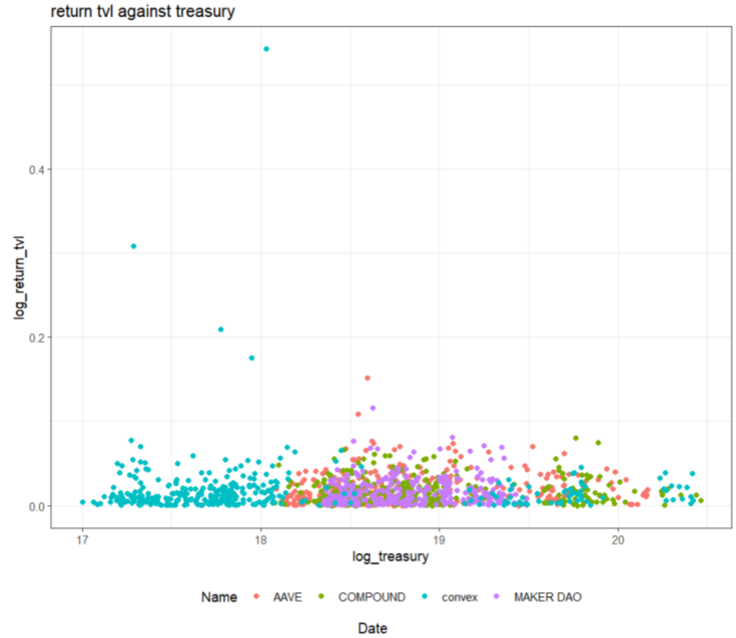

Figure 8 Dispersion of data observed between tvl and Treasury (cross-sectional approach)

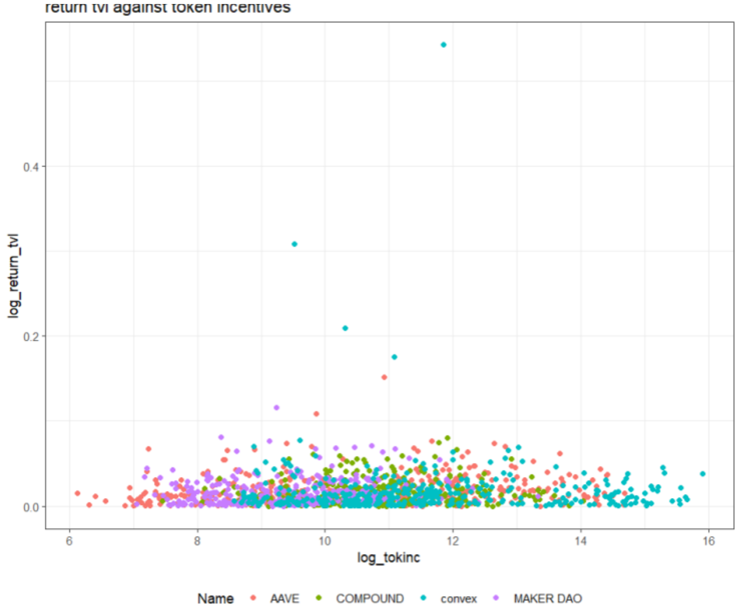

Figure 9 Dispersion of data observed between tvl and token-based incentives (cross-sectional approach)

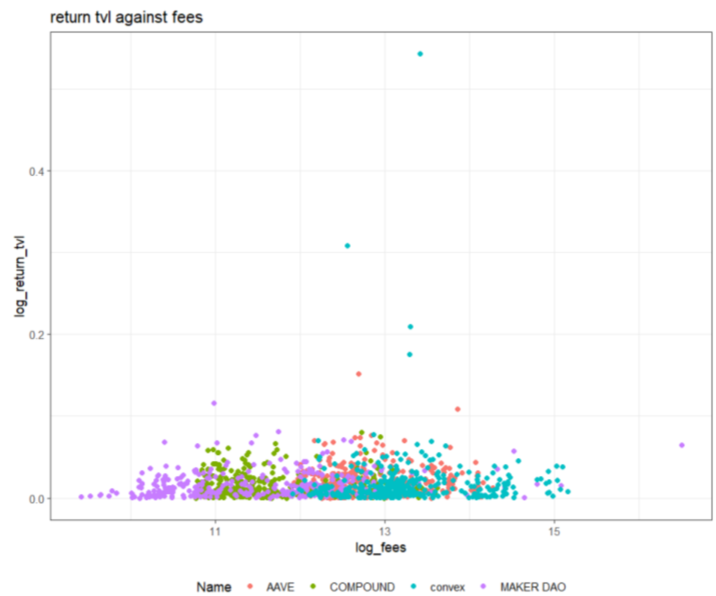

Figure 10 Dispersion of Observed Data Between TVL and Treasury (Cross-Sectional Approach)

Those graphs show high dispersion among observable variables. We notice a cluster of data point on the lower part of each graph indicating none-linear relationship between the dependent and independent variable. This relationship may vary across different protocols. This plot might suggest heterogeneity across group.

It would be useful to explore fixed-effects or random-effects models in the analysis to account for the differences between these groups.

Stationarity test

We apply the **Levin-Lin-Chu (LLC) test** that is a statistical method used to test for stationarity in panel data where you have multiple cross-sectional units:

- Null hypothesis (H0): The time series contains a unit root and is non-stationary (if the p-value > 0.05)
- Alternative hypothesis (H1): The time series contains no unit root and is considered stationary (if the p-value ≤ 0.05)

Table 3: Levin-Lin-Chu (LLC) test

Variables	normalized t-statistic	P-value	Stationarity
Log Return TVL	-2.49	0.006	Stationary
Log TVL	-2.81	0.002	Stationary
Log Token Incentives	-3.49	0.000	Stationary
Log Treasury	-1.74	0.04	Stationary
Log Fees	-1.34	0.09	Not Stationary
Log Core Dev	**-1.17**	**0.121**	**Not Stationary**

Here the LLC test is used to check for unit roots in a panel dataset. It is commonly applied to determine whether the time series in a panel are stationary or not. The column "normalized t-statistic" and "p-value" are use determine the stationarity of each variable.

The LLC test might not be sufficient to assess the stationarity of the variables. The **Augmented Dickey-Fuller (ADF) test** is then also needed to test the stationarity among time series data.

- Null hypothesis (H0): The time series contains a unit root and is non-stationary (if the p-value > 0.05)

- Alternative hypothesis (H1): The time series contains no unit root and is considered stationary (if the p-value ≤ 0.05)

Table 4: Augmented Dickey-Fuller (ADF) test

Variables	Dickey-Fuller Statistic	P-value	Stationarity
Log Return TVL	-5.76	0.01	Stationary
Log TVL	-3.36	0.059	Not Stationary
Log Token Incentives	-2.88	0.203	Not Stationary
Log Treasury	-3.21	0.086	Not Stationary
Log Fees	-2.02	0.566	Not Stationary
Log Core Dev	-1.46	0.803	Not Stationary

Lag order: 7

The Augmented Dickey-Fuller (ADF) test was performed to check for stationarity in the time series data. The null hypothesis of the test states that the series has a unit root (non-stationary), while the alternative hypothesis suggests stationarity. Variables with p-values below 0.05 are stationary, rejecting the null hypothesis, while those with p-values above 0.05 indicate non-stationarity. Based on the results, only "log_return_tvl" is stationary. These results inform decisions on whether transformations are needed for further analysis

None of the independent variables are stationary. Initially, the stationarity of the dependent variable, "log TVL," was uncertain. However, it was later confirmed to be non-stationary, justifying the use of its return as the dependent variable. Consequently, each independent variable is different to ensure stationarity.

F-test for fixed effect model or pooled mode:

To test if the fixed effects (individual-specific or time-specific) significantly improve the model fit compared to a pooled OLS model, which assumes no differences between cross-sectional units or time periods. We will run a **F-test.**

- **Null Hypothesis (H_0):** Fixed effects are not required; the coefficients of individual (or temporal) effects are equal to zero, indicating that a simple OLS model is sufficient.
- **Alternative Hypothesis (H_1):** Fixed effects are required; significant differences exist between entities (or time periods).

Table 5: F-test

Model	F-statistic	Degree of freedom (df1) (df2)	P-value	Alternative Hypothesis
Pooled Regression	**1.48**	**(12) (384)**	**0.128**	**Unstability**

Here, the F-test allow to test whether a pooled regression model (i.e., a model that assumes all cross-sectional units share the same coefficients) is appropriate, or whether there are significant differences across groups that would justify using a fixed effects model.). The "F-statistic and "p-value" are used to determine which model between the fixed and OLS is best suited for our data.

H0 is not rejected and thus the pooled OLS model can be used. However, as we specified earlier in the methodology and seen on the chart, we preferred fixed or random effect.

Hausman test for fixed effect model or random effect:

The **Hausman test** is used to determine whether a fixed effects model (FEM) or a random effects model (REM) is appropriate by testing for correlation between the individual-specific effects (αi) and the explanatory variables (Xit).

- **Null Hypothesis (H_0):** cov(αi, Xit) = 0 (if p-value > 0,05, the REM is appropriate)
- **Alternative Hypothesis (H_1):** cov(αi, Xit) ≠ 0 (if p-value ≤ 0,05, the FEM is appropriate)
-

Table 6: Hausman test

Test	Chi-squared	Degree of freedom df	P-value	Alternative Hypothesis
Hausman	**0.566**	**4**	**0.966**	**One model is inconsistent**

Here, the Hausman test allow to check if individual attributes (Xit) appear to have no correlation with the unobserved individual effects (αi). The "Chi-quared" and "p-value" are used to determine which model between the fixed and random effect is best suited for our data .

We fail to reject H0 and thus we use the random effect model.

$$\text{log\_return\_tvl}_{it} = \alpha + \beta_1 \cdot \text{diff\_log\_fees}_{it} + \beta_2 \cdot \text{diff\_log\_tokinc}_{it} + \beta_3 \cdot \text{diff\_log\_coredev}_{it} + \beta_4 \cdot \text{diff\_log\_treasury}_{it} + \mu_i + \lambda_t + \epsilon_{it}$$

Table 7: result model regression (random effect)

Variables	Coefficient	Standard Error	Z-value	P-value	Significance
Intercept	0.000868	0.00006	13.58	2.2*e-16	Significant
diff log fees	0.00017	0.00044	2.6	0.0091	Significant
diff log Tokinc	0.000053	0.00004	1.36	0.1726	Not Significant
diff log Core Dev	0.000332	0.00028	1.15	0.2483	Not Significant
diff log Treasury	0.000842	0.00029	2.86	0.0042	Significant

Here are the results of the regression with random effect applied (without correction). The "coefficients" represent the value of the estimate from the regression (in other word the beta). The "p-value" and "z-value" are used to determine if each estimator are not biased.

Table 8: Check for potential bias with the error term

Test	Statistic	P-value	Conclusion
Durbin-Watson	1.96	0.39	No serial correlation
Breush-Pagan	15.25	0.004	Heteroscedasticity present
Breush Pagan LM	140.98	< 2.2*e-16	Cross-sectional dependence present
Pesaran CD	10.58	< 2.2*e-16	Cross-sectional dependence present

Here is a table showing potential bias from the random effect regression. The column "statistic" and "P-value" are use to determine if there are presence of specific bias with the error term.

Verification of the model's significance:

Other tests were conducted to verify the significance of the random effects model. These tests revealed positive serial correlation (Durbin-Watson test), positive heteroscedasticity (Breusch-Pagan test), and positive cross-sectional dependence (Pesaran CD test).

Therefore, adjustments are needed to correct these biases and improve the model's significance:

In order to enhance the accuracy of our model and mitigate potential bias, we employed clustered standard errors. The advantage of using clustered standard errors lies in their ability to account for serial correlation within the error terms of each entity, a feature that standard heteroskedasticity-robust errors fail to address.

Table 9: Results of the Model Coefficients after Adjustment using random effect (final result):

Variables	**Estimate**	**Standard Error**	**Z-value**	**P-value**	**Significance**
Intercept	0.000868	0.00006	13.08	< 2.2*e-16	Significant
diff log fees	0.00017	0.00002	5.74	1.7*e-8	Significant
diff log Tokinc	0.000053	0.00001	2.97	0.003	Significant
diff log Core Dev	0.000332	0.00045	0.73	0.461	Not Significant
diff log Treasury	0.000842	0.00033	2.5	0.012	Significant
R – squared	***0.0475***			***AIC***	***-5289.17***
Adjusted R-squared	***0.039***			***BIC***	***-5268.60***

The following presents the results of the regression model with random effects, adjusted for clustering standard errors. The "coefficients" represent the estimated values (i.e., the β coefficients) from the regression. The "p-value" and "z-value" are used to assess the statistical significance of each estimator, determining whether they are unbiased and significantly different from zero. Additionally, the Information Criteria (AIC and BIC) and R-squared provide insights into the model's fit, evaluating how well the model explains the variation in the dependent variable and the overall explanatory power of the independent variables.

Explanation and discussion:

The results derive from a random effects regression model aims to explain the drivers of the TVL into a protocol.

The low values of the information criteria (AIC: -5289 and BIC: -5268) suggest that the model provides a good fit of the observed data, avoiding overfitting and does not include unnecessary variables.

Here are the interpretations of our independent variables:

- "Fees": a 1% change in fees is associated with a 0.0117% change in the dependent variable, all else being equal. This aligns with the hypothesis that higher fees are correlated with increased value locked in the protocol, reflecting higher user engagement or willingness to pay.
- "Token incentive": a 1% change in token incentives corresponds to a 0.0053% change in the dependent variable. This supports the idea that higher token rewards attract more participants, thereby increasing the protocol's performance or TVL.
- "Core developer": The number of core developers does not have a statistically significant effect on the dependent variable. This result suggests that the variation in developer count may not directly influence the TVL.
- "Treasury": Larger treasuries are viewed as a sign of financial health and stability, which positively impacts the protocol's performance. A 1% increase in treasury size leads to a 0.0843% increase in the dependent variable.

Around 5.93% of the variation in the dependent variable is explained by the independent variables. The adjusted R-squared, at 4.984%, suggests that after adjusting for the number of predictors in the model, the explained variance remains modest.

This low explanatory power implies that other factors, not captured by the model, likely play a significant role in determining the dependent variable. This metric is not significant in time series, especially when using the return as dependent variable.

The R-squared measures the proportion of variance in the dependent variable explained by the model, relative to the variance of the dependent variable around its mean. However, when using return, the mean of returns is often remarkably close to zero, especially for financial time series. As a result, the variance around the mean might be very small, and thus the R-square value may be misleading.

Limitations of the Model and Potential Improvements:

This study is limited by the small sample size, focusing on only four protocols (AAVE, MakerDAO, Compound, and Convex).
Expanding the dataset to include more protocols, especially from diverse markets, would improve the generalizability of the results.

While the model includes key variables such as fees, treasury size, and token incentives, it omits key factors as: average loan amounts, average loan duration, collateral utilization rates; or macroeconomic variables like market volatility and regulatory announcements.

These could provide a broader perspective on what drives TVL. Some variables, such as developer activity, were found to be insignificant, suggesting the need to refine variable selection and explore potential multicollinearity.

Finally, introducing interaction terms, non-linear relationships, could reveal deeper insights and address model limitations.

Monte Carlo Simulation

This study evaluates the robustness of the relationship between explanatory variables (Fees, Price, Operating Expenses, Token Incentives, and Treasury) and the dependent variable, log Return_TVL, using Monte Carlo simulations. By introducing controlled variations in the explanatory variables, we assess the stability and variability of the regression coefficients, providing insights into the model's sensitivity to data fluctuations.

We conducted 10,000 simulations to get a solid and reliable sample analysis. The average of the simulated coefficients represents their central tendency, while the standard deviation quantifies their dispersion. These two parameters helped us to understand both the typical outcome and the range of variability in the simulation.

The graphs and table below summarize the simulation results, showing the distribution of coefficient estimates. Each histogram represents a variable, with the x-axis indicating estimates and the y-axis their frequency. This analysis highlights the bias, variance, and reliability of the estimators under simulated conditions.

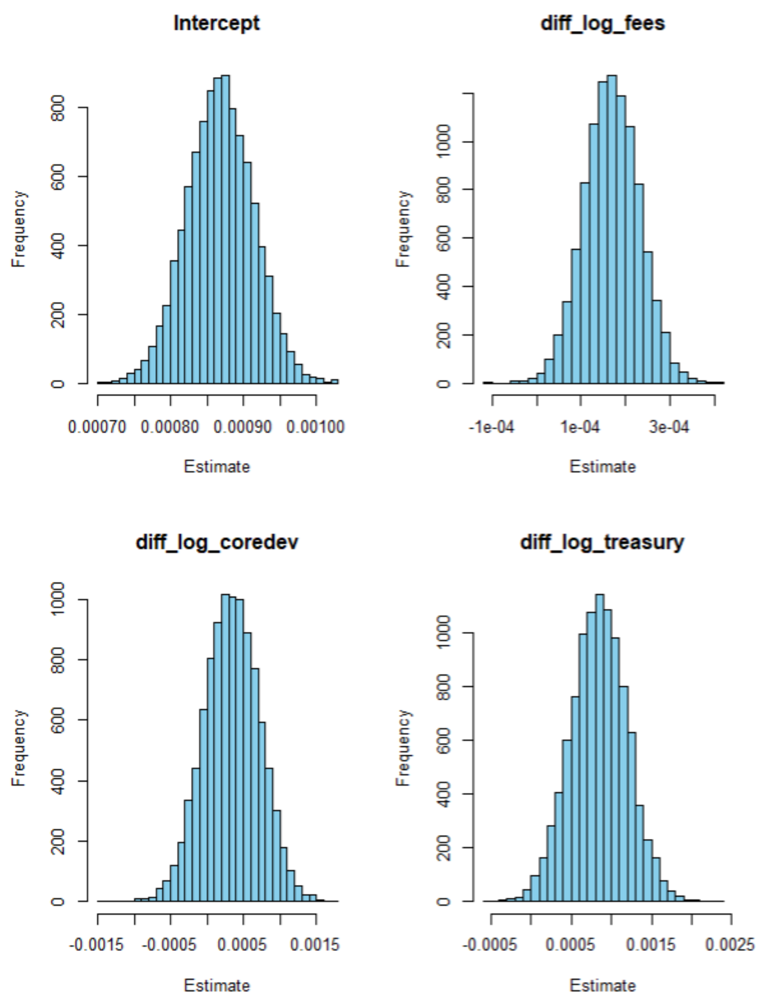

Figure 11 Distribution of Simulated Coefficient Estimates from Monte Carlo Simulation

Explanation of the Results:

The histograms exhibit symmetric and approximately bell-shaped distributions for all estimated coefficients, suggesting that the estimators are unbiased and follow a normal distribution under the given simulation setup. Below, we provide an interpretation of the observed results.

- **Intercept**: The distribution of the intercept estimates is centered around 0.0009, with the frequency peaking at approximately 800 occurrences. This indicates that the intercept estimate is consistently close to this value across simulations, demonstrating a high degree of precision.

- **diff_log_fees**: The histogram for diff_log_fees is centered near 0, with most estimates falling within a small range. The high peak frequency of about 1,250 suggests strong precision, while the symmetry indicates minimal bias in the estimator.

- **diff_log_tokinc**: The estimates for diff_log_tokinc also form a symmetric and tight distribution, implying that the model provides reliable and consistent estimates for this variable. The frequency distribution suggests limited variability in estimates.

- **diff_log_coredev** and **diff_log_treasury**: These coefficients show slightly wider distributions compared to others, implying greater variability in the estimates. However, the distributions remain symmetric and centered, signifying that the estimators are still unbiased, albeit with slightly less precision than the other ones.

Symmetry and Normality of Distributions:

The histograms reveal that all coefficient estimates are **symmetric and bell-shaped**, indicating that the estimators follow a normal distribution (as expected for OLS estimators under standard assumptions). This symmetry implies:

- The regression estimators are **unbiased** (on average, the estimates center around the true value).
- The **central limit theorem** applies effectively, as the distributions do not show skewness or heavy tails.

Overall, the Monte Carlo simulation confirms that the regression model's estimators are unbiased and consistent, with varying levels of precision depending on the variable.

The results are also summarized in the below table:

Table 10: Monte Carlo Simulation table

Variables	True estimates	Simulated Mean	Difference from mean (%)	Simulated Median	Difference from median (%)
Intercept	0.000868	8.68 e-04	0.011	8.68 e-04	0.00144
diff log fees	0.00017	1.69 e-04	-0.18	1.68 e-04	-0.624
diff log Tokinc	0.000053	5.29 e-05	-0.16	5.25 e-05	-0.893
diff log Core Dev	0.000332	3.38 e-04	1.94	3.38 e-04	1.908
diff log Treasury	0.000842	8.44 e-04	0.25	8.47 e-04	0.629

Assessing Bias in Coefficient Estimates:

The summary table compares the **true coefficients** to the **simulated mean and median coefficients** across all simulations. The percentage differences help us measure **bias**, which occurs when the simulated estimates deviate systematically from the true values.

Minimal Bias:

- The **Intercept** has the smallest percentage bias (0.011% for the mean and 0.001% for the median), indicating that it is well-estimated with negligible bias.
- **Fees** and **Tokinc** exhibit slightly larger negative biases (e.g., -0.19% and -0.16%, respectively). These values are still within acceptable limits, meaning the estimates are close to the true values.
- **CoreDev** shows the largest positive bias (~1.94% difference from the mean). This suggests a tendency for the model to slightly overestimate the impact of **CoreDev** under the simulated conditions. While this bias is small, it could be meaningful in sensitive applications or with larger magnitudes of the coefficient.
- Regarding **Treasury**, although the bias percentage of 0.25% is higher than that of Intercept, Fees, and Tokinc, it still demonstrates that the simulation is highly accurate and that the model used produces reliable results.

The simulation confirms that the regression model's estimators are **unbiased overall**, with all coefficients staying close to their true values. However, small biases in certain variables like **CoreDev** might warrant closer investigation, particularly if this variable plays a critical role in the analysis. Futhermore, In the empirical test this was also the only non-significative variable.

Conclusion

This study provides a comprehensive analysis of the factors influencing the performance of decentralized finance lending and borrowing protocols, focusing on TVL as a key indicator of success. Through panel regression and Monte Carlo simulations, we evaluated the relationships between explanatory variables that are fees, token incentives, treasury size, core developer activity, and TVL returns.

The results demonstrate that certain variables such as fees, token incentives, and treasury size significantly impact TVL, highlighting their importance in driving user engagement and protocol stability. While core developer activity showed no significant influence, this might reflect indirect or context-dependent effects. The Monte Carlo simulation further confirmed the strength of our regression estimators, revealing minimal bias and consistent reliability across most variables.

However, the study also uncovered limitations, such as the restricted dataset and low explanatory power of the model. Expanding the sample size, incorporating additional variables, and exploring non-linear relationships could provide more nuanced insights. Addressing these limitations and refining the model would enhance its predictive accuracy and applicability.

Overall, this research contributes valuable knowledge to the emerging field of DeFi, offering a foundation for future studies to build on and providing practical implications for improving the sustainability and growth of lending and borrowing protocols in decentralized finance.

The Future of DeFi Lending and Borrowing: Unlocking Global Financial Potential

Decentralized Finance (DeFi) is rapidly transforming the financial landscape, with lending and borrowing protocols at the upfront of this revolution. As blockchain technology continues to mature, DeFi promises to democratize access to financial services, transcending geographical, economic, and institutional barriers. The future of DeFi lending and borrowing lies in enhanced scalability, integration with traditional finance, and the adoption of innovative mechanisms like real-world asset tokenization and under-collateralized loans. By leveraging smart contracts, artificial intelligence, and more secure infrastructure, DeFi can address current challenges such as volatility, regulatory uncertainties, and user experience limitations. Eventually, these advancements have the potential to redefine global finance, making capital more accessible, transparent, and efficient for individuals and businesses worldwide.

Sources

Alexis Collomb & klara Sok , 2016, *Blockchain / Distributed Ledger Technology (DLT): what impact on the Financial Sector ?* [link]

Arianna Agosto & Alessia Cafferata, 2020, *Financial Bubbles: A study of Co-Explosivity in the cryptocurrency Market* [link]

Biplab Kumar and Inder Sekhar Yadav, 2018, *financial development and economic growth: panel evidence from BRICS [link]*

Bartoletti, M., Chiang, J.H.Y., and Lluch-Lafuente A., (2020). *SoK: Lending Pools in Decentralized Finance. [link]*

Betul Kaplan , Vahit Ferhan Benli, Elcin Aykac Alp, December 15 2023, *Istanbul Finance Congress, Decentralized Finance and new protocols [link]*

Corbet, Shaen, John W. Goodell, Samet Gunay, and Kerem Kaskaloglu. 2021. *Are DeFi Tokens a Separate Asset Class from Conventional Cryptocurrencies?* SSRN Electronic Journal: 3810599. [link]

Dominik Metelski,Janusz Sobieraj,2022, *"Valuations of decentralised finance (DeFi) protocols" Int. J. Financial Stud. 2022, 10(4)*

Fadel Dabaja, Tobias Dahlberg,2021, *"Decentralized Finance and the Crypto Market: Indicators and Correlations"*

Gilles Brice M'bakob, june 2024, University of Douala, Faculty of Economics and Applied Management Sciences, Cameroon , *Bubbles in Bitcoin and Ethereum: The role of halving in the formation of super cycles / sustainable futures, volume 7,* [link]

Greene W. (2018), PART IV, *Econometric Analysis, eighth edition, Pearson - Prentice Hil*

jiahoua XU & Nikhil Vadgama, 11 January 2022, *From Banks to DeFi: the Evolution of the Lending Market*

Michael Sockin,Wei Xiong, 2022, *decentralization through tokenization,national bureau of economic research paper n°29720.*

MakerDAO. (n.d.). Getting started: Maker Protocol 101. [link]

Popescu, Andrei-Dragoş. 2020. *Decentralized Finance (DeFi)-The lego of finance. Social Sciences and Education Research Review* 1: 321–49.

Panel Data Analysis Fixed and Random Effects using Stata (v. 6.0) Oscar Torres-Reyna [link]

Șoiman, Florentina, Guillaume Dumas, and Sonia Jimenez-Garces, 2022, *The return of (I) DeFiX. arXiv arXiv:2204.00251.* [link]

Shay Palachy Affek, 2019, *stationarity in time serie* [link]

Will Schmitt, New York, July 23 2024, *Financial times, The SEC's approval of ether ETFs give crypto a firmer toehold in traditional US finance* [Link]

GITHUB FILE [link]

CHAINANALYSIS [link]

TOKENTERMINAL [link]