Financial Innovation

RESEARCH

Open Access

Mega-whale driven decentralized lending protocols

Xiaotong Sun[1*], Charalampos Stasinakis[2] and Georgios Sermpinis[2]

*Correspondence:
Xiaotong.Sun@waikato.ac.nz

[1] Waikato Management School, University of Waikato, Hamilton 3240, New Zealand
[2] Adam Smith Business School, University of Glasgow, Gilbert Scott Building, Glasgow G12 8QQ, UK

Abstract

This study examines systemic risks in decentralized lending protocols stemming from "mega-whales"—large users with transaction volumes exceeding $100 million—who dominate liquidity dynamics. Using high-frequency transaction-level data from major protocols (*Aave v2/v3*, *Compound v2/v3*, *Morpho Blue*, *Flux*, and *SparkLend*) spanning August 2024 to August 2025, we identify 550 mega-whales and analyze their impacts across market size, structural risks, and user activity. Key findings reveal the following: (1) The net growth in both mega-whale deposits and borrowing is negatively correlated with the market-size dynamics of lending protocols, while arbitrage-driven behaviors (such as token diversification and platform switching) further fragment the growth of Total Value Locked (TVL); (2) Although whale borrowing serves to temporarily stabilize leverage ratios, deposit activities may adversely impact systemic stability; however, an increase in mega-whale deposits is found to positively stimulate the growth and participation of unique depositors; (3) Utilizing event-window controls around DeFi shocks, we confirm that these effects are strongly associated with mega-whale activities rather than market-wide events. These results point to participation patterns that may be inconsistent with DeFi's egalitarian ethos, suggesting that economic concentration could influence market dynamics.

Keywords: Lending protocols, Decentralized finance, Mega-whale users

Introduction

In May 2022, the Terra ecosystem's collapse erased over $40 billion in value, triggered in part by manipulative borrowing and price inflation by large actors, underscoring how "whales" can destabilize decentralized finance (DeFi) systems (Chiu et al. 2023). While DeFi lending protocols (LPs) promise transparent, intermediary-free peer-to-peer borrowing and lending via blockchain and smart contracts—reducing costs and enhancing financial inclusion (An and Rau 2019; Liu et al. 2020; Harvey et al. 2021; Ali 2024)—such incidents highlight potential vulnerabilities in DeFi markets, motivating further examination of how concentrated user behavior may influence liquidity and participation dynamics (Gogel et al. 2021).

Prior research has identified multifaceted risks in DeFi, including smart contract vulnerabilities, high volatility and leverage-induced financial instability, decentralized governance challenges, regulatory uncertainties, and interconnected systemic threats

(Capponi et al. 2023; Weingärtner et al. 2023). A critical yet understudied dimension is user concentration, where "whales"—large cryptocurrency holders—exhibit distinct, often more profitable behaviors than retail investors, amplifying market inefficiencies (Chernoff and Jagtiani 2025). In lending contexts, features such as anonymity, overcollateralization, and rigid terms exacerbate fragility through price-liquidity feedback loops, which whales exploit (Chiu et al. 2023). Empirical evidence from protocols such as Compound and Aave shows extreme concentration: the top 100 addresses control 75% of deposits and 78% of loans, far surpassing traditional banking levels (Saengchote 2023). Whales often pursue speculative yield farming or governance influence, heightening risks (Capponi et al. 2023).

Although studies have modeled DeFi lending economics and systemic risks (Gudgeon et al. 2020b; Bartoletti et al. 2021; Castro-Iragorri et al. 2021), large-scale empirical analyses of "mega-whales"—users with volumes over \$100 million[1]—remain scarce, particularly regarding their behavioral impacts on market dynamics, risks, and participation (Sun et al. 2024). This gap is pressing, as whale dominance challenges DeFi's narrative of broad-based, egalitarian participation, even within technically decentralized systems, by fostering an illusion of decentralization (Sun et al. 2024).[2]

This paper investigates how mega-whale behavior correlates with liquidity, market size dynamics, and user participation across lending protocols. While these patterns may relate to broader stability concerns discussed in prior research, our analysis focuses on empirical associations rather than direct causal estimation of systemic risk. To do this, we use transaction-level data from *Aave v2/v3*, *Compound v2/v3*, *Morpho Blue*, *Flux*, and *SparkLend* (August 13, 2024, to August 13, 2025), identifying 550 such users. We examine effects across market size (e.g., TVL, volumes), structural risks (e.g., leverage ratios), and user activity (e.g., unique participants). We test three hypotheses: (H1) While mega-whale deposit and borrowing positions are negatively correlated with protocol growth, their transaction frequency significantly expands market size; (H2) Mega-whale activities exert divergent effects on structural stability, where borrowing serves as a temporary stabilizer for leverage ratios, whereas concentrated deposits tend to increase fragmentation and risk exposure; (H3) Beyond aggregate volumes, mega-whale deposit variations and speculative behaviors (such as platform switching) positively influence the growth of unique depositors. The findings indicate that while mega-whale transaction frequency supports overall market volume, their directional deposit and borrowing positions are negatively correlated with protocol growth. Regarding structural risks, empirical evidence suggests that mega-whale borrowing serves as a temporary stabilizer for leverage ratios, whereas concentrated deposits tend to increase fragmentation and risk exposure. Furthermore, variations in mega-whale deposits and platform-switching behaviors are found to positively stimulate user growth.

[1] The \$100 million transaction volume threshold for defining "mega-whales" is selected based on prior academic research, such as Chernoff and Jagtiani (2025). In their study, crypto "whales" are defined as entities with substantial holdings or transaction volumes (implicitly exceeding \$100 million, drawing from analyzes of Ethereum and Bitcoin wallets that control 1–5% of network activity). This threshold is appropriate, as it identifies actors whose activities are associated with price manipulations and liquidity disruptions, consistent with the emphasis on concealed effects in Ethereum-based lending protocols. In the supplementary material, we do robustness checks by using \$200 million as an alternative threshold, the results show consistency.

[2] The presence of whale users causes a range of centralization concerns within decentralized systems. For a more comprehensive discussion, we direct readers to Sai et al. (2021).

Our contributions are threefold. Theoretically, we provide empirical evidence on concentration in decentralized markets, extending banking theories of depositor/borrower discipline (Berger and Turk-Ariss 2014; Santos and Winton 2019) to DeFi, indicating parallels with liquidity dynamics discussed in the literature. However, we do not directly estimate systemic risk, and our interpretations should be understood within this empirical scope. Empirically, we offer granular insights via high-frequency data and event studies, bridging gaps in the whale behavior literature (Saengchote 2023; Chernoff and Jagtiani 2025). Practically, the results inform protocol designers (e.g., whale monitoring tools) and regulators (e.g., mitigating risks under frameworks such as Ali (2024) and Weingärtner et al. (2023)), advancing the innovation-regulation debate.[3]

The paper proceeds as follows: Sect. "Background and hypothesis development" provides the LP background and hypothesis development; Sect. "Whale users in DeFi lending" characterizes mega-whales; Sect. "How mega-whales drive lending protocols" presents the empirical results; and Sect. "Conclusion" concludes. The appendix and supplementary material include technical details.

Background and hypothesis development

Lending protocols (LPs)

Based on blockchain technology, LPs resemble banks in crypto markets, allowing their users to borrow and lend cryptocurrencies (Bartoletti et al. 2021). Any agent can lend their cryptocurrency to an LP. Similar to depositors in banks, LP depositors can earn interest by providing liquidity. For any cryptocurrency, all deposits will be stored in a lending pool, which can be borrowed by anyone. To initiate loans, borrowers should first lock collateral, usually cryptocurrencies accepted by LPs, and both loans and interests should be repaid if a borrower aims to unlock their collateral assets.

Figure 1 illustrates how borrowers and depositors interact with LPs using DAI stablecoin and Aave as an example. If an LP supports a loan of some cryptocurrency, a liquidity pool will be generated, where users can deposit or borrow the cryptocurrency. Panel B presents the balance sheet view of DAI stablecoin in the LP, addressing the importance of sufficient liquidity.

In addition to borrowers and depositors, liquidators also play an important role in LPs. When a borrower fails to repay their loan, liquidators can (partly) repay the failed loan. As a result, liquidators can purchase the borrower's collateral at a discount. Usually, the process is defined as liquidation. Compared to traditional bank lending, a pivotal difference of LPs is that key parameters of loans are not decided by third parties. For example, LPs apply different mathematical models to determine interest rates. Other suggested changes, e.g., new acceptable collateral, will usually be jointly decided by LP users by voting. For more details, we refer readers to the summary research presented by Werner et al. (2022) and Gudgeon et al. (2020a).

Currently, diversified LPs coexist in DeFi, while several mainstream LPs account for most lending activities. In this paper, we focus on several leading LPs, including *Aave v2/v3*, *Compound v2/v3*, *Morpho Blue*, *Flux*, and *SparkLend*.

[3] Although our primary hypothesis emphasizes the driving influence of mega-whales on protocol liquidity, the empirical evidence also highlights an endogenous feedback loop between these actors and the protocols.

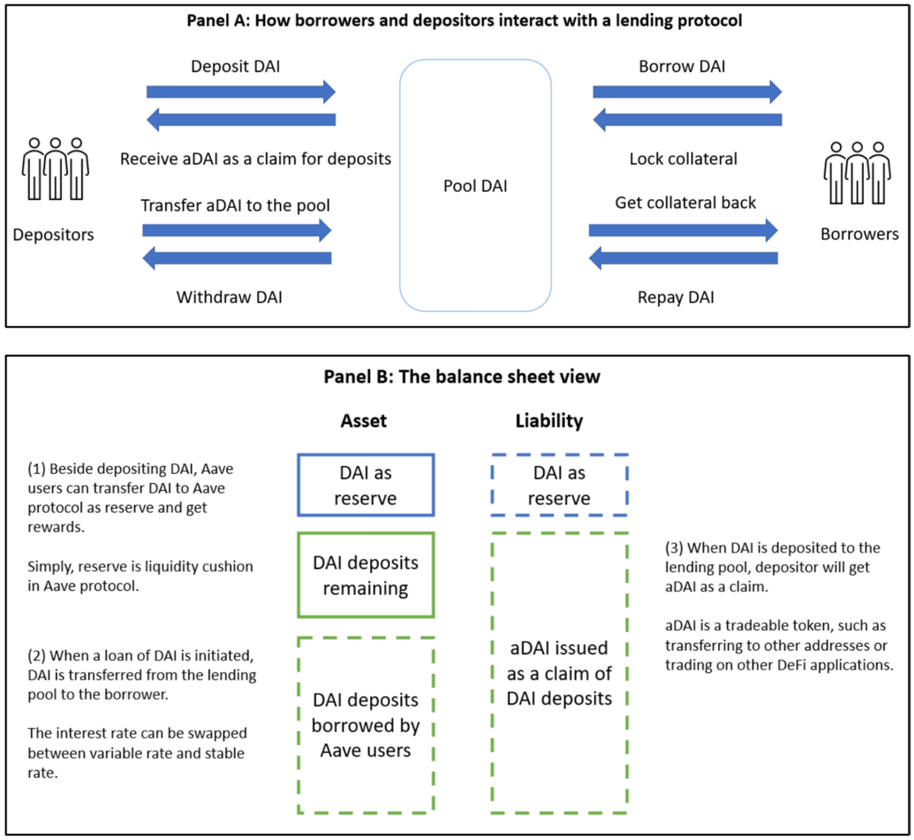

Fig. 1 Pooled funds in lending protocols. *Note* This figure illustrates borrowing and lending in LPs by using Aave as an example. For each token, there will be a pool. Panel A shows how users interact with the lending pool of DAI. Depositors can deposit their token (i.e., DAI) and receive an amount of claim (i.e., aDAI). When depositors want to withdraw their tokens, they need to transfer the claim to the LP. Borrowers need to lock collateral when requiring loans. When they successfully repay loans, collateral can be returned. Panel B presents the balance sheet view of the lending pool of DAI. Both DAI deposits and DAI reserves are classified as assets. DAI loans represent the debt obligations of borrowers. When an Aave user deposits DAI, the aDAI token will be issued as a claim; therefore, aDAI is a part of liability. More details about reserve mechanism in Aave can be found: https://docs.aave.com/aavenomics/incentives-policy-and-aave-reserve

In addition to significant market capitalization, the leading LPs expand their influence by introducing new features. For example, Aave was the first LP to introduce 'flash loans'. Put simply, flash loans do not require any collateral since the loan will be borrowed and repaid in an atomic transaction group (Qin et al. 2021b). More details on the use of flash loans can be found in Gudgeon et al. (2020a), Qin et al. (2021a), and Wang et al. (2021). This function allows for more on-chain transactions. Aave introduced several innovative features in V2,[4] such as swapping collateral assets and repaying debts with collateral assets. These innovative technical features not only increased the adoption of Aave but also made it the main industry standard.

[4] https://github.com/aave/protocol-v2

Hypothesis development

Transaction frequency and market size dynamics

The literature exhibits a distinct duality regarding how mega-whale users influence the market size of decentralized lending protocols. On the one hand, high-frequency trading and cross-platform arbitrage activities are often viewed as drivers of market growth. Studies by Heimbach and Huang (2024) and Cornelli et al. (2025) find that frequent interactions by large users expand liquidity pools and the total value locked (TVL) of protocols. This perspective is further supported by McLaughlin et al. (2023) in their study of the Ethereum arbitrage ecosystem, suggesting that high-frequency arbitrage transactions significantly boost market activity. From this viewpoint, mega-whales—through agile liquidity migration such as platform switching—act as market integrators that bridge fragmented liquidity and provide essential support for protocol scaling.

On the other hand, when focusing on the net volume of deposits or borrowings by mega-whales, the literature reveals significant negative risks. The concentration of large-scale capital often triggers a crowding-out effect and undermines the long-term robustness of the protocol. First, information asymmetry and rent extraction (Azar et al. 2024) enable mega-whales to leverage private information for arbitrage gains; such behavior increases transaction costs for average users and leads to market fragmentation. Second, concentrated deposit and borrowing activities may trigger adverse selection exploitation. According to the loss-versus-rebalancing (LVR) theory proposed by Milionis et al. (2025), professional arbitrageurs (mega-whales) systematically extract value from passive liquidity providers (retail users). This increased cost can drive retail users to withdraw their liquidity, thereby inhibiting the sustained expansion of the market size.

Furthermore, governance risks and the concentration of market power are critical factors limiting growth. Han et al. (2025) point out that mega-whales may utilize concentrated voting power to pursue short-term private interests rather than the overall value of the protocol, a governance risk that erodes user confidence. This systemic instability caused by large-holder behavior often manifests as abrupt TVL drawdowns during extreme market conditions. In traditional finance theory, Pritsker (2002) similarly argued that the position adjustments of large investors carry a substantial price impact, absorbing liquidity that would otherwise serve the broader market, thus creating a crowding-out effect.

In summary, the impact of mega-whale behavior on market size dynamics is not unidirectional but depends on specific behavioral patterns. While frequent arbitrage and switching may enhance efficiency, concentrated positions may suppress growth through risk amplification and cost transference. Based on this, we propose the following hypothesis:

H_1: Mega-whale activities exhibit a dualistic effect on market size dynamics: their directional deposit and borrowing positions are significantly and negatively correlated with protocol liquidity growth, whereas their speculative cross-platform activity (e.g., platform switch ratio) positively contributes to market expansion.

Influences on liquidity dynamics and structural risks

Mega-whales play a dual role in lending protocols: they act as both indispensable liquidity providers and potential sources of systemic risk. The extant literature suggests that mega-whales significantly enhance liquidity efficiency by providing large-scale deposits, which in turn improves borrowing accessibility and reduces transaction slippage (Kaplan et al. 2023). However, such high capital concentration introduces profound structural vulnerabilities.

From the supply side, mega-whale deposit activities are characterized by "search-for-yield" motives and high capital mobility. According to the theoretical framework of Cornelli et al. (2025), institutional-grade depositors in DeFi are acutely sensitive to cross-platform interest rate differentials and protocol security. When mega-whales engage in "platform switching" to capture higher yields or mitigate perceived risks, these large-scale capital withdrawals directly trigger volatility in TVL and lead to liquidity fragmentation. Chiu et al. (2023) further model this fragility, demonstrating that concentrated withdrawals by large holders in stressed environments can activate a "price-liquidity feedback loop." In this cycle, depleting liquidity and falling collateral prices reinforce each other, exacerbating the protocol's systemic risk exposure.

In contrast, mega-whale borrowing activities theoretically exhibit greater "strategic inertia." Unlike deposit behaviors driven by short-term interest rate arbitrage, borrowing is typically associated with long-term leverage strategies, hedging requirements, or the acquisition of protocol governance rights. Aramonte et al. (2022) note that borrowing positions held by large participants involve higher switching costs and strategic considerations, thus appearing more stable than deposit positions during market turbulence. By maintaining consistent utilization rates, such behavior objectively serves to balance the protocol's balance sheet and stabilize structural risk ratios. Saengchote (2023) also discusses how stable borrowing demand can mitigate fragmentation risks by optimizing resource allocation.

Furthermore, the concentration of governance power among mega-whales may influence structural resilience from an institutional perspective. Studies on protocols such as Compound and MakerDAO reveal that whale-centric governance models, while improving decision-making efficiency in the short term, can exacerbate structural risks through centralized power (Papangelou et al. 2023; Sun et al. 2024). Consequently, the impact of mega-whales on protocol stability likely diverges between their depositing and borrowing functions. Based on these theoretical insights, we propose the following hypothesis:

H_2: Mega-whale borrowing activities will enhance liquidity efficiency and stabilize structural risk profiles (e.g., via higher deposit-to-borrow ratios) in LPs, while their deposit activities will increase fragmentation and risk exposure (e.g., lower deposit share of volume).

Moderation of user activity metrics

The literature suggests that the participation of large-scale actors significantly shapes the entry and engagement patterns of other market participants. Following the logic that "liquidity begets liquidity," mega-whale deposits provide the essential depth required to minimize transaction slippage and optimize lending rates, creating a conducive environment that attracts retail users (Gudgeon et al. 2020a; Werner et al. 2022). Within the information-asymmetric environment of DeFi, these large-scale capital injections serve as a potent "quality signal," where retail investors interpret whale deposits as an endorsement of the protocol's security and yield stability (Cornelli et al. 2025).

This dynamic is further reinforced by "rational herding" behavior in lending markets, where subsequent participants observe the actions of sophisticated lead investors to mitigate perceived risks (Zhang and Liu 2012). Furthermore, the speculative cross-protocol activities of mega-whales—often captured by platform switching—act as a beacon for arbitrageurs and yield seekers. By navigating between protocols to capture interest rate differentials, mega-whales signal profitable opportunities, thereby stimulating broader transaction volumes and user activity. Consequently, we argue that mega-whale activities function as a catalyst for user growth rather than a deterrent in the growth phase of protocols. Based on these theoretical insights, we propose our third hypothesis:

H_3: Increases in mega-whale deposits will positively influence depositor growth in lending protocols, and their speculative activities (e.g., cross-protocol platform switching) will positively attract and stimulate broader user engagement.

While the whale driven hypothesis emphasizes the impact of mega-whales on protocol liquidity and risk dynamics, it is essential to acknowledge that these actors do not operate in a vacuum. The relationship between mega-whales and decentralized protocols is likely reflexive (Soros 1987). Specifically, protocol-level incentive mechanisms—such as fluctuating interest rates, borrow utilization, and liquidity rewards—serve as critical signals that attract or repel mega-whales.

This potential for an endogenous feedback loop implies that while whales influence the protocol's state, the protocol's state simultaneously shapes whale behavior. Consequently, our empirical strategy, particularly the use of Granger causality tests, is designed to identify the predictive lead-lag relationship (i.e., which variable moves first in time) rather than claiming a strictly unilateral structural causality. By doing so, we account for the complex interplay inherent in DeFi ecosystems.

Whale users in DeFi lending

Data source

To obtain a comprehensive understanding of the decentralized lending market, we focus on several mainstream LPs, including *Aave v2/v3*, *Compound v2/v3*, *Morpho Blue*, *Flux*, and *SparkLend*. This paper investigates how mega-whale behavior correlates with liquidity and market size dynamics.[5] Because major lending protocols (e.g., Aave, Compound, Morpho) deploy smart contracts separately on different blockchains, each

[5] While prior DeFi research often examines individual liquidity pools (e.g., Saengchote 2023), our choice of the protocol–chain unit is primarily necessitated by data constraints, as high-frequency, granular data for individual pools are not consistently available across all protocols and chains in our sample. Notwithstanding these practical limitations, this level of aggregation is also conceptually appropriate for our study: mega-whale activity frequently spans multiple assets and pools within the same protocol through multiasset strategies and rapid redeployments. Consequently, the protocol–chain level captures the relevant economic environment for liquidity and systemic risk transmission without artificially fragmenting whale behavior that is strategically executed at the protocol scale.

protocol–chain combination operates as an independent liquidity environment with distinct asset compositions, utilization dynamics, interest-rate parameters, and liquidation rules. Within each chain, we aggregate across token-specific pools to construct a chain-level panel for each protocol. This approach preserves economically meaningful variation across chains while avoiding noise from illiquid or sparsely used pools (e.g., tail-asset pools with negligible volume). The resulting dataset reflects the operative unit at which liquidity, leverage, and risk evolve in practice. To ensure the timeliness of the research findings, we analyze the period from August 13, 2024, to August 13, 2025. This timeframe is strategically chosen for two reasons:

Market Dynamics Post-COVID: The cryptocurrency market exhibited distinct characteristics before, during, and after the COVID-19 pandemic, with these differences exerting significant influence on the decentralized lending market. For instance, systemic fragility in decentralized markets during COVID-19 volatility is documented in Lehar and Parlour (2022), which highlights how liquidity crises in protocols such as Compound and Aave were exacerbated by arbitrageur behavior and price impacts.

Regulatory Shifts: Rapid regulatory changes since mid-2024 have reshaped the DeFi landscape. For example, the U.S. Congress repealed tax reporting requirements for noncustodial DeFi brokers in April 2025 under the Congressional Review Act, reversing prior regulations. Simultaneously, the EU's 6th Anti-Money Laundering Directive (6AMLD) and California's Digital Financial Assets Law (effective July 2026) reflect heightened scrutiny of crypto activities. These policies directly influence user preferences for regulated centralized platforms over DeFi, as analyzed in Auer and Claessens (2018) and Feinstein and Werbach (2021).

To retrieve the datasets, we utilize Flipside Crypto[6] as our data source. Flipside Crypto serves as a specialized blockchain analytics platform that provides structured, SQL-ready data for academic and industry research. The platform aggregates on-chain transaction histories decoded smart contract events and wallet activity across 35 + blockchains, including Ethereum, Solana, and Bitcoin, enabling granular analysis of decentralized finance (DeFi) protocols and user behavior. Its data architecture is designed for interoperability, with consistent schemas across chains to facilitate cross-protocol comparisons. Table 1 and Figs. 2, 3 below present an overview of the studied LPs in our analysis.

Whale users in decentralized lending

After obtaining granular datasets (at the transaction level) for the decentralized lending market, we are able to identify whale users. By considering the total borrowing and lending volumes on prominent LPs across various blockchains, we can filter out the largest users.

There are 550 mega-whales (with a total trading volume exceeding USD 100 million) and 386 super whales (with a total trading volume ranging from USD 50 million to USD 100 million). Table 2 provides an overview of the activities of mega-whales and super whales. Although the number of these two types of users is relatively small compared to the total user base of LPs, their contribution to the total volume is substantial. The

[6] https://flipsidecrypto.xyz/home

Table 1 Overview of the lending protocols

Protocol	Blockchain	Total lending volume (in $m)	Total borrowing volume (in $m)	Total unique lenders daily sum	Total unique borrowers daily sum	Avg daily utilization ratio
aave v2	ethereum	28,329.17	2598.161	13719	4220	0.1236
aave v2	polygon	1426.033	20.63315	326101	1038	0.022316
aave v2	avalanche	88.59978	1.810468	22762	496	0.032427
aave v3	ethereum	921729.3	153466.3	269156	186460	0.184376
aave v3	polygon	76164.04	2203.424	341247	129661	0.423826
aave v3	base	66194.22	110859.2	876664	282893	1.454939
aave v3	arbitrum	16443.31	6080.041	615765	302359	0.382248
aave v3	avalanche	7735.592	3960.831	154881	68101	0.464589
aave v3	optimism	5371.504	7063.618	233316	93412	1.119262
aave v3	gnosis	452.161	197.2081	34445	9457	0.498913
compound v2	ethereum	1480.231	389.5724	4108	2749	1.210225
compound v3	ethereum	4777.859	8982.625	13537	28411	4.014998
compound v3	arbitrum	1327.462	3965.138	24254	45309	7.63158
compound v3	base	275.7813	619.4021	5638	112045	33.57779
compound v3	optimism	147.0238	566.773	20964	36489	15.77033
compound v3	polygon	87.86395	182729.2	8422	40916	25080696
flux	ethereum	70.97927	31.39082	155	10	0.069224
morpho blue	base	292778.3	240140.8	286959	45402	0.137595
morpho blue	ethereum	105948.7	15968.56	67321	33940	0.271222
spark	ethereum	27714.66	10552.97	10133	9412	0.661836
spark	gnosis	31.23252	15.58663	1801	1530	57.44593

This table provides a comprehensive summary of the lending and borrowing activities undertaken by prominent LPs over the period from August 13, 2024, to August 13, 2025. For each LP, we compiled key performance indicators across various blockchains, acknowledging that leading LPs typically operate on multiple mainstream blockchains. *The total lending volume* and *total borrowing volume*, both expressed in millions of USD, offer a broad indication of the market acceptance of these LPs. *The total unique lenders (borrowers) daily sum* for each LP represents the sum of daily distinct participants, serving as an indicator of user engagement. Additionally, *the average daily utilization ratio*, calculated as the mean of the daily borrowings in USD divided by the daily deposits in USD over the sample year, provides insight into the potential liquidity risks associated with each LP

trading strategies of these users are likely to vary as well. We find that some whale users do not engage in any borrow/deposit transactions. Therefore, some whales focus more on borrowing cryptocurrencies, while in the transaction records of other whales, the deposit volume accounts for a larger proportion. Some whales have executed approximately 30,000 transactions, whereas the average number of transactions in the sample year is less than 1,000. Consequently, there must be some whales that are high-frequency traders, while others do not exhibit such high trading frequency. These disparities suggest that these users have their own unique trading strategies. As a result, we have become interested in how the transaction patterns of these users affect LPs. Given that the trading volume of super whales is relatively small compared to that of mega-whales, we will focus solely on mega-whales in the following sections.

To gain deeper insight into the trading strategies employed by mega-whales, we constructed a series of variables derived from their transaction records. Specifically, we have retained only those transactions with a transaction volume exceeding $10,000. The definitions of these variables are outlined in Table 3.

Net deposit and *net borrow* metrics directly quantify mega-whales' roles as liquidity providers or borrowers within LPs. A positive net deposit indicates capital inflow,

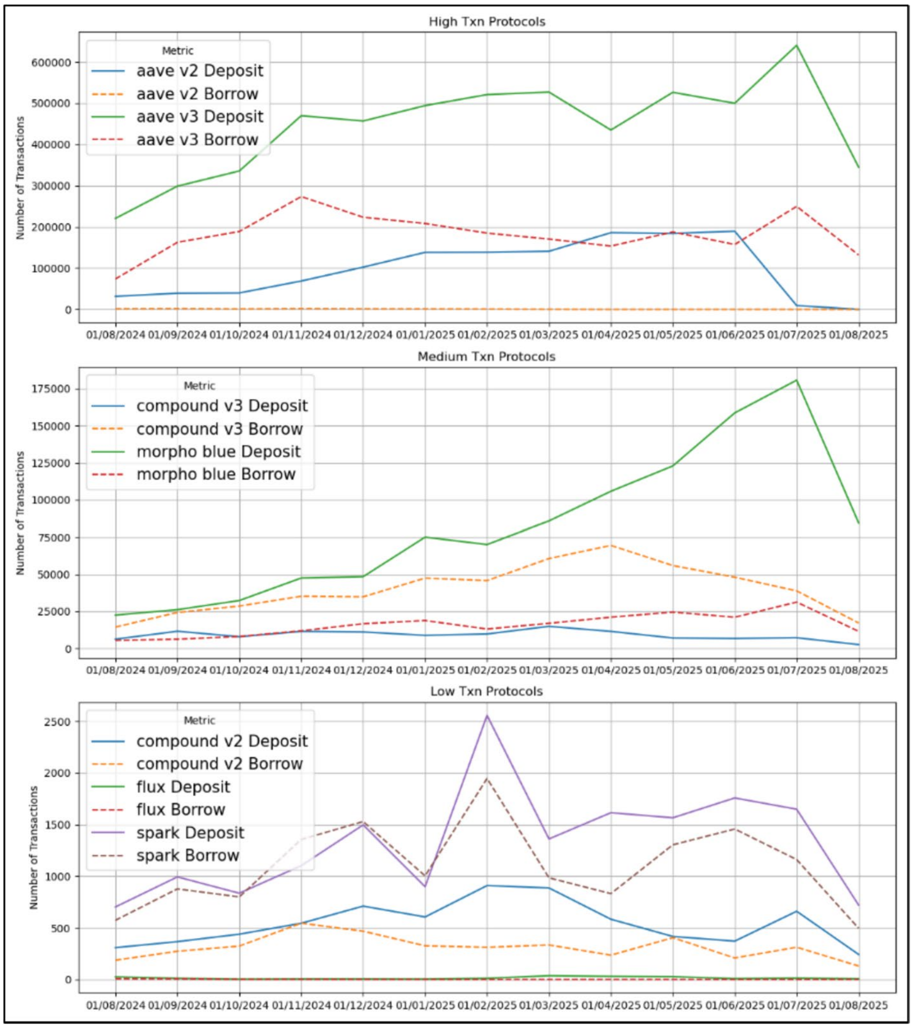

Fig. 2 Transaction counts in lending protocols. *Note* This figure illustrates the monthly transaction counts across major decentralized lending protocols from August 2024 to August 2025. The data are categorized into two transaction types: deposit and borrow transactions, with protocols grouped by transaction counts (*High Txn*, *Medium Txn*, and *Low Txn*). The vertical axis employs a logarithmic scale to enhance readability of magnitude variations

enhancing protocol solvency and interest rate stability. Conversely, sustained net borrowing may strain liquidity pools, particularly if borrow utilization (defined below) keeps increasing, potentially triggering liquidation cascades or interest rate spikes. These metrics thus serve as leading indicators of protocol-level risk exposure.

The *Txn count* reflects the frequency of mega-whale interactions with protocols. A high *txn count* suggests active portfolio rebalancing or arbitrage activities, which may amplify volatility in utilization rates. For instance, frequent deposits/withdrawals could destabilize interest rate equilibrium, while rapid borrowing/repayment cycles may create short-term liquidity shocks.

A high *token count* implies diversified asset holdings, potentially indicating speculative strategies such as cross-token arbitrage or yield farming across protocols. Similarly,

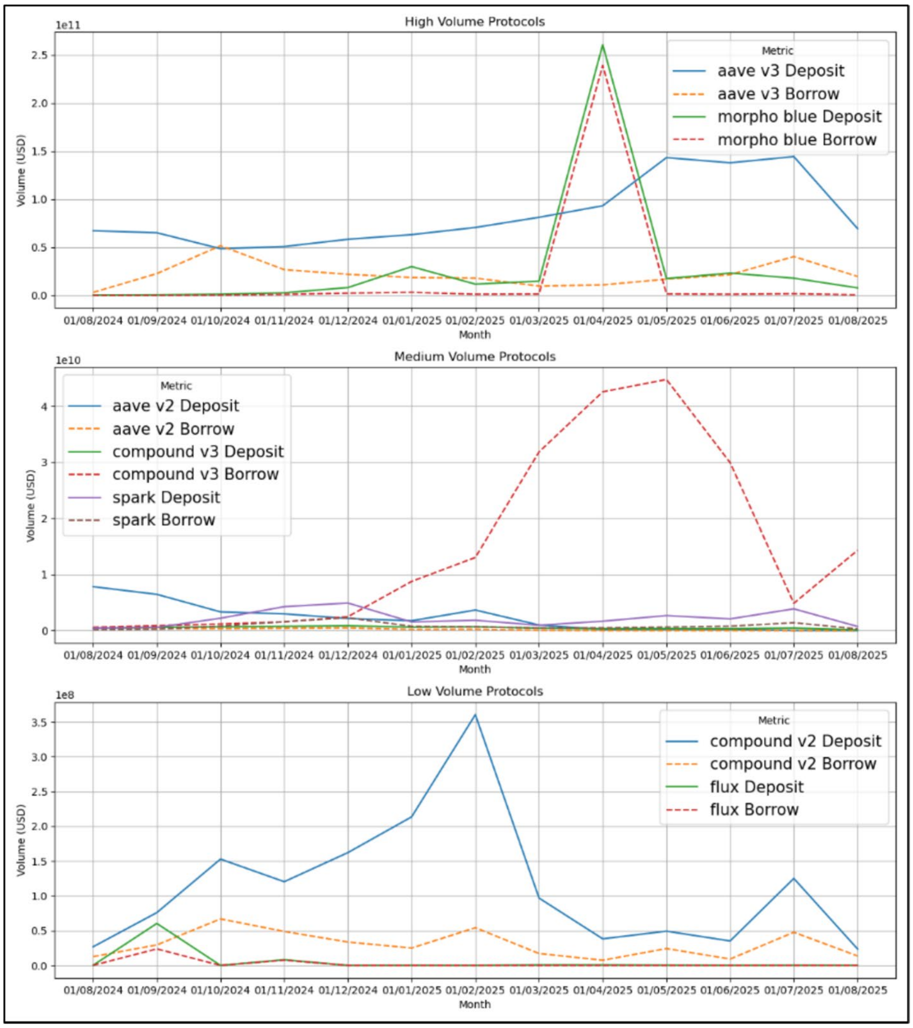

Fig. 3 Lending and borrowing volumes in lending protocols. *Note* This figure illustrates the monthly trading volume across major decentralized lending protocols from August 2024 to August 2025. The data are categorized into two transaction types: deposit and borrow volumes, with protocols grouped by transaction volume (*High Volume*, *Medium Volume*, and *Low Volume*). The vertical axis employs a logarithmic scale to enhance readability of magnitude variations

an elevated *platform switch ratio* signals opportunistic behavior, where mega-whales exploit price/yield differentials across decentralized finance (DeFi) platforms. These activities may enhance market efficiency but also introduce systemic risk through correlated withdrawal patterns or concentrated liquidity extraction.

The *large trade ratio* (transactions ≥ $1 million) is particularly impactful due to mega-whales' capacity to alter deposit/loan composition. A single large deposit can suppress borrowing costs, while massive withdrawals may trigger liquidity crashes. Moreover, market participants often monitor mega-whale transactions as signals, leading to herding behaviors that amplify price/yield movements. This creates a feedback loop where mega-whale actions indirectly dictate broader market sentiment.

Borrow utilization (borrowed/deposited assets) is a critical leverage indicator. High values suggest unsustainable borrowing and increasing liquidation risks during adverse

Table 2 Descriptive statistics of mega-whales and super whales

	Total volume (in $ million)	Deposit volume (in $ million)	Borrow volume (in $ million)	Total txns	Protocol used
Panel A: Mega-whales					
Mean	2127.15	1843.51	283.64	623.67	1.91
Std	11,038.47	10,463.21	1165.16	2492.09	1.08
Min	100.07	0	0	1	1
25%	144.20	95.02	17.16	33	1
50%	257.57	165.42	71.29	102.5	1
75%	643.13	447.64	157.84	269.25	2
Max	168,353.55	155,796.68	14,320.82	29,386	6
N	550	550	550	550	550
	Total volume (in $ million)	**Deposit volume (in $ million)**	**Borrow volume (in $ million)**	**Total txns**	**Protocol used**
Panel B: Super whales					
Mean	70.23	45.19	25.04	128.06	1.66
Std	13.97	20.29	19.06	198.96	0.81
Min	50.14	0	0	1	1
25%	57.62	31.64	8.01	22	1
50%	69.35	45.95	24.97	52	1
75%	81.10	56.95	36.07	125.5	2
Max	99.96	97.93	88.94	1393	5
N	386	386	386	386	386

This table provides a comprehensive summary of the lending and borrowing activities undertaken by mega-whales and super whales in prominent LPs over the period from August 13, 2024, to August 13, 2025

Table 3 Characteristics-related variables of mega-whales

Variable	Definitions
Net deposit	The daily net amount of deposited cryptocurrencies (in $ million) across major LPs by mega-whales
Net borrow	The daily net amount of borrowed cryptocurrencies (in $ million) across major LPs by mega-whales
Txn count	The daily count of lending and borrowing transactions conducted by mega-whales
Token count	The daily count of distinct cryptocurrencies traded by mega-whales
Platform switch ratio	The ratio of the number of times mega-whales switch across different LPs to the total number of transactions, calculated on a daily basis
Large trade ratio	The ratio of the number of large transactions (with a volume of at least $1 million) to the total number of transactions, calculated on a daily basis
Borrow utilization	The ratio of the daily amount of borrowed cryptocurrencies (in $ million) to the daily amount of deposited cryptocurrencies (in $ million). It is noteworthy that this ratio can exceed 1

This table introduces several variables about mega-whales' characteristics

price movements. As mega-whales dominate protocol TVL, their utilization shifts directly affect protocol stability. For example, a sudden drop in borrow utilization may signal deleveraging, prompting retail users to withdraw deposits and exacerbating liquidity shortages.

Figure 4 shows how mega-whales trading patterns change over the sample year. The variables exhibit significant volatility, with noticeable spikes occurring at specific dates.

These fluctuations suggest that the trading patterns of mega-whales are subject to considerable variation across different time periods. Such variability may reflect responses to market events, shifts in strategy, or changes in risk appetite, underscoring the dynamic nature of large-scale trading activities in the cryptocurrency market.

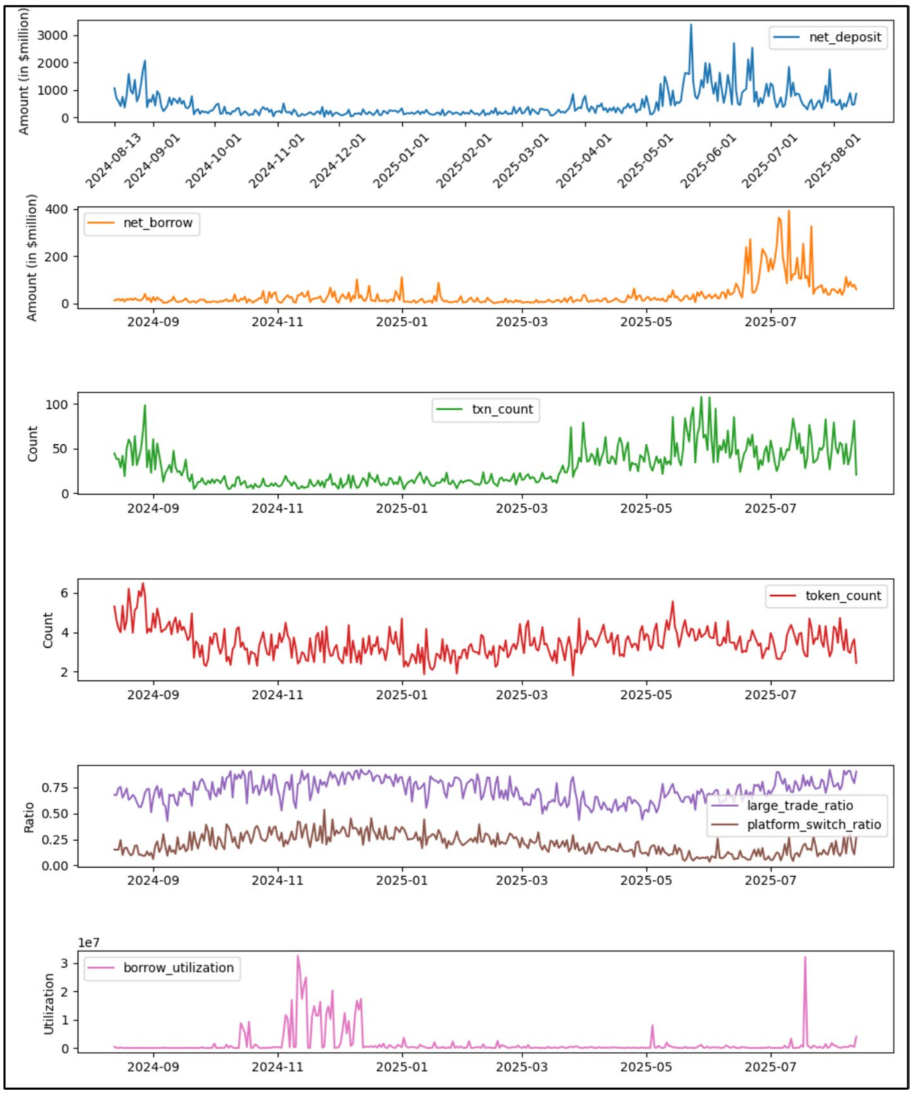

Fig. 4 Characteristics of mega-whales. *Note* This figure shows the characteristics-related variables of mega-whales over the period from August 13, 2024, to August 13, 2025. Variables are computed daily across all 550 mega-whales and studied protocols (Aave v2/v3, Compound v2/v3, Morpho Blue, Flux, SparkLend). Summed metrics—net deposit (total USD inflows minus outflows), net borrow (total USD borrowings minus repayments), and txn count (total transactions)—are aggregated as sums to reflect collective activity. The token count represents the unique number of tokens traded daily across the group. Ratios—platform switch ratio (proportion of transactions involving protocol switches), large trade ratio (proportion of trades ≥ $1 million), and borrow utilization (borrowed/deposited assets)—are calculated as simple averages across active mega-whales to capture typical behavior. Data are raw and not normalized, sourced from Flipside Crypto. The data exhibit significant volatility, with noticeable spikes occurring at specific dates

Different types of whale users in decentralized lending

A key challenge in DeFi is that wallet addresses do not reveal investor identities. We therefore cannot directly label mega-whales as institutions versus individuals, nor can we observe their off-chain objectives. Nevertheless, the transaction-level traces recorded on-chain allow us to form economically grounded expectations about the strategies that these large participants are likely to employ. Such a strategy-based characterization is useful for two reasons. First, it clarifies how the same "whale activity" can map into different economic mechanisms in lending protocols. Second, it provides a structured interpretation of the volatility observed in mega-whale characteristics—such as fluctuations in token count, transaction frequency, and platform switching—over the sample period.

We construct a set of mega-whale characteristics from transaction records, capturing their roles as liquidity providers versus borrowers and their intensity and complexity of interaction with lending protocols. In particular, we use *daily net deposit* and *net borrow* to quantify whether whales act primarily as liquidity providers or liquidity demanders. *Txn count* measures the frequency of whale interactions with protocols, while *token count* captures the breadth of assets involved in whale activity. The *platform switch ratio* captures cross-protocol mobility (i.e., how often whales switch across different lending protocols relative to the total number of transactions). A *large trade ratio* measures the share of very large transactions, reflecting whales' capacity to alter pool composition through discrete large flows. *Borrow utilization*, defined as borrowed assets relative to deposited assets, summarizes aggregate leverage pressure and can exceed one.

These characteristics have direct economic interpretations. A higher *txn count* typically indicates active portfolio rebalancing, rapid borrowing/repayment cycles, or opportunistic execution that can transmit short-term liquidity shocks. Higher *token count* and *platform switch ratio* are consistent with cross-token arbitrage, yield/collateral optimization, and rate-shopping across protocols. A *large trade ratio* captures the discreteness of capital flows: a small number of large deposits or withdrawals can move interest rate conditions and liquidity availability even if the number of transactions is limited.

To operationalize investor characterization without relying on unobservable identity labels, we adopt a behavioral clustering approach. Using the full transaction histories of the 550 mega-whales, we aggregate wallet-level features that correspond to (i) deposit-versus-borrow orientation (e.g., borrow share based on total borrowed and deposited amounts), (ii) execution intensity (transactions per active day), (iii) cross-protocol mobility (platform count and platform switch ratio), (iv) asset complexity (token count), and (v) discreteness of flows (large trade ratio). We log-transform highly skewed variables, standardize all inputs, and apply an unsupervised clustering algorithm (K-means) to obtain four interpretable groups. Importantly, the goal is not to "identify" who these investors are but rather to organize heterogeneous on-chain behavior into economically meaningful strategy archetypes.

The resulting clusters map naturally into four strategy types.

- *Cross-protocol reallocators/arbitrage-motivated switchers (cluster 1)* Whales in this group frequently operate across multiple lending protocols and switch platforms relatively often. Their activity involves a broader set of tokens, consistent with opportunistic reallocation when yield, borrowing costs, or collateral efficiency changes across protocols. Economically, these whales are most aligned with a "capital mobility" channel. This would mean that rather than leaving assets as sticky deposits, they rotate positions across protocols and tokens, which can enlarge market activity while contributing to fragmentation of liquidity.

- *Execution-heavy/high-throughput whales (cluster 2)* This cluster is characterized by extremely high interaction intensity (very frequent transactions per active day) but comparatively low platform switching. Such patterns are consistent with whales that repeatedly adjust leverage, collateral, and borrow positions within a protocol (or a small set of protocols), potentially reflecting systematic rebalancing, sophisticated execution strategies, or rapid responses to market movements. This type is most relevant for mechanisms where whale activity intensity amplifies short-run changes in protocol-level liquidity conditions and transaction volume.
- *Episodic large movers/passive-treasury style whales (cluster 3)* Whales in this cluster exhibit very low borrowing shares, use few tokens and typically a single protocol, and conduct only a handful of transactions. However, their trades are disproportionately large, indicating infrequent but consequential capital flows (e.g., one-time deposits, withdrawals, or treasury reallocations). Economically, this type corresponds to "lumpy" liquidity provision/withdrawal. This means that even when transaction counts are small, discrete large flows can substantially affect pool conditions, interest rates, and perceived stability.
- *Single-protocol leveraged position managers (cluster 4)* This group shows substantial borrowing activity but little to no cross-protocol switching, indicating a preference for managing leveraged positions within a particular lending venue. Such behavior is consistent with a leverage/hedging or collateral optimization strategy anchored in one protocol. This type is most directly tied to leverage-related channels (e.g., borrow utilization dynamics) and to how concentrated borrowing demand may transmit risk under adverse price movements.

Table 4 reports the characteristics of the median whales in each of the four behavioral clusters. For each cluster, we present median values of whale-level measures that capture borrowing orientation, transaction intensity, asset complexity, cross-protocol mobility, and the discreteness of capital flows. Using medians mitigates the influence of extreme outliers and highlights the typical behavior within each cluster.

This typology clarifies how the same observable variables can carry different meanings depending on the dominant strategies in the whale population on a given day. For example, increases in token count and platform switch ratio are more naturally interpreted as shifts toward reallocator/arbitrage behavior rather than passive yield provision. Similarly, spikes in txn count are most consistent with execution-heavy whales engaging in rapid rebalancing or borrowing/repayment cycles, whereas a high large trade ratio with low txn count is consistent with episodic large movers whose impact operates through discrete flows. By providing these plausible behavioral expectations, the typology improves the interpretability of the patterns in whale characteristics and offers a coherent bridge from whale behavior to the hypotheses developed in Sect. "Background and hypothesis

Table 4 Median whale characteristics across four clusters

Cluster	borrow_share	txn_per_active_day	txn_count	token_count
Panel A: Borrowing orientation, activity intensity, and asset complexity				
1	0.086737	15.63636	1356	9
2	0.404602	2.821549	139	7
3	0.004201	2	5.5	2
4	0.358587	2.796825	68	4
Cluster	**platform_count**	**platform_switch_ratio**	**large_trade_ratio_p95**	**span_days**
Panel B: Cross-protocol mobility, trade discreteness, and activity span				
1	2	0.00919	0.017442	267.4115
2	3	0.196938	0	287.2497
3	1	0	0.666667	6.310347
4	1	0	0	195.8097

This table reports the median whale-level characteristics within each cluster. Clusters are obtained using an unsupervised clustering procedure (K-means) applied to wallet-level aggregates constructed from transaction records. *Borrow_share* is defined as the ratio of total borrowed USD value to the sum of total borrowed and deposited USD value for a whale. *Txn_per_active_day* is the average number of whale transactions per active day, and *Txn_count* is the total number of whale transactions in the sample. *Token_count* denotes the number of distinct tokens involved in the whale's lending-protocol interactions. *Platform_count* is the number of distinct lending protocols used by the whale. *Platform_switch_ratio* measures how frequently the whale switches across protocols relative to its total number of transactions. *Large_trade_ratio_p95* is the fraction of the whale's transactions whose USD exceeds the 95th percentile of the full transaction-size distribution (computed over all whale transactions), capturing the discreteness of large capital movements. *Span_days* is the number of days between the whale's first and last observed transactions in the sample

development" and the protocol-level outcomes examined in Sect. "How mega-whales drive lending protocols".

How mega-whales drive lending protocols

In this section, we will examine how mega-whales influence the decentralized lending market. We will first introduce the key metrics that can measure the status of LPs and then present the empirical results in the following subsections.

Variable definition

We select several variables to capture the operational dynamics, financial resilience, and network effects inherent in decentralized lending markets. This framework integrates three complementary analytical dimensions: market size dynamics, structural risk indicators, and user activity, each contributing to a holistic understanding of protocol viability. The table below introduces the variable definitions.

From a financial intermediation perspective, market size metrics (Panel A of Table 5) serve as primary liquidity indicators. The 1-day log-differences in TVL, borrowed assets, and deposited capital ($\Delta ln\ TVL$, $\Delta ln\ Total\ Borrow$, $\Delta ln\ Total\ Deposit$) provide granular insights into short-term capital reallocation behaviors. These metrics operate as leading indicators for market sentiment shifts, with TVL changes being particularly sensitive to macroeconomic conditions and protocol-specific governance events. The inclusion of trading volume dynamics ($\Delta ln\ Total\ Volume$) further enriches this dimension by capturing transactional intensity, which often correlates with arbitrage activities and speculative trading positions.

Table 5 Key metrics of the decentralized lending market

Factor	Definition
Panel A: Market size dynamics	
Δln TVL	Daily log-differences of total value (in $ million) locked in the prominent LPs
Δln Total borrow	Daily log-differences of total amount (in $ million) of borrowed cryptocurrencies in the prominent LPs
Δln Total deposit	Daily log-differences of total amount (in $ million) of deposited cryptocurrencies in the prominent LPs
Δln Total volume	Daily log-differences of total trading volume (in $ million) of the prominent LPs
Panel B: Structural risk indicators	
ΔDeposit share of volume	The first difference of the ratio between the total amount (in USD) of deposited cryptocurrencies and the total trading volume (in USD) of prominent LPs
ΔDeposit to borrow ratio	The first difference of the ratio between the total amount (in USD) of deposited cryptocurrencies and the total amount (in USD) of borrowed cryptocurrencies
Panel C: User activity	
Δln Unique depositors	Daily log-differences of number of distinct depositors in the prominent LPs
Δln Unique borrowers	Daily log-differences of number of distinct borrowers in the prominent LPs
Δln Total unique users	Daily log-differences of daily number of distinct users in the prominent LPs
Δln Txn count protocol	Daily log-differences of total number of transactions in the prominent LPs

This table introduces a series of factors related to LPs. The factors can be retrieved from *flipsidecrypto.xyz*

Decentralized lending protocols feature multiple dimensions of risk, including collateral-price volatility, collateral concentration, liquidation risk, and cross-pool liquidity imbalances. While several alternative measures (such as volatility of collateral assets, collateral diversification indices, or relative reserve levels across pools) can potentially capture aspects of protocol fragility, these variables are difficult to compare consistently across chains, tokens, and protocol designs. In contrast, *the deposit share of volume* and the *deposit-to-borrow ratio* offer (i) cross-protocol comparability, (ii) daily observability, and (iii) direct links to solvency and liquidity conditions within lending pools.

Panel B of Table 5 introduces the structural risk indicators.[7] Specifically, the *deposit-to-borrow ratio* is a core solvency metric in both theoretical and empirical analyses of DeFi lending (e.g., Gudgeon et al. 2020b; Capponi et al. 2023), as undercollateralization and liquidity mismatches trigger liquidation cascades. Similarly, the *deposit share of volume* captures liquidity concentration and the extent to which inflows are driven by deposit-heavy mega-whales rather than balanced borrowing activity, which is a direct channel through which large users can fragment liquidity or destabilize utilization dynamics. These indicators are therefore well suited for cross-protocol, chain-level analysis and provide a transparent and tractable way to capture structural risk without imposing strong assumptions about collateral types or valuation models.

User activity indicators (Panel C of Table 5) introduce behavioral finance elements into the analysis. Metrics tracking daily log-differences in unique depositors/borrowers and daily active users (*Δln Unique Depositors*, *Δln Unique Borrowers*, *Δln Total Unique Users*) reflect the network adoption of LPs. Transaction count volatility (*Δln Txn Count*

[7] To acknowledge broader dimensions of risk, we additionally discuss in Sect. "Structural risk indicators" the roles of collateral volatility, collateral diversification, and cross-pool liquidity imbalances as alternative metrics, and highlight how future work may integrate those measures once high-frequency, pool-specific collateral composition data are more systematically available across protocols and chains.

protocol) serves as a proxy for operational friction, with declining activity often preceding platform migration to competing protocols.

To summarize, this variable selection strategy adheres to three methodological principles: (1) multidimensionality ensuring coverage of liquidity, risk, and behavioral dimensions; (2) temporal sensitivity through 1-day change metrics capturing market responsiveness; and (3) risk-adjusted perspective by incorporating both nominal changes and ratio-based stress indicators. The framework provides a robust foundation for the empirical analysis of decentralized lending ecosystems. Table 6 presents the descriptive statistics of these variables.

Market size dynamics

We start our analysis by investigating the influence of mega-whales on the market size dynamics of LPs. In addition to independent variables relevant to mega-whales, we also integrate various categories of control variables. First, we introduce several variables

Table 6 Descriptive statistics of dependent variables

	Δln TVL	**Δln Total borrow**	**Δln Total deposit**	**Δln Total volume**
Panel A: Market size dynamics				
Mean	0.02	−0.01	0.01	0.01
Std	1.64	1.38	1.60	1.46
Min	−5.91	−14.65	−9.56	−9.63
25%	−0.85	−0.48	−0.54	−0.43
50%	0.02	−0.01	−0.01	−0.01
75%	0.81	0.52	0.56	0.46
Max	11.47	9.47	13.59	14.26
	ΔDeposit share of volume	**ΔDeposit to borrow ratio**		
Panel B: Structural risk indicators				
Mean	5549.05	0.00		
Std	561,262.94	0.17		
Min	−17040591.43	−1.00		
25%	−1.16	−0.05		
50%	0.00	0.00		
75%	1.29	0.04		
Max	17532584.19	0.97		
	Δln Unique depositors	**Δln Unique borrowers**	**Δln Total unique users**	**Δln Txn count protocol**
Panel C: User activity				
Mean	0.00	0.00	0.00	0.00
Std	0.40	0.40	0.36	0.41
Min	−4.44	−2.40	−4.44	−5.25
25%	−0.17	−0.18	−0.15	−0.19
50%	0.00	0.00	0.00	−0.01
75%	0.16	0.17	0.14	0.18
Max	2.86	2.56	2.59	2.76

This table reports descriptive statistics of the dependent variables. To ensure full alignment with the econometric specifications, all dependent variables are reported in their exact estimation scales. Specifically, market-size and user-activity outcomes are presented as daily log-differences, while structural risk indicators are presented as first differences. Monetary values are expressed in millions of USD where applicable. The data source is retrieved from *flipsidecrypto.xyz* associated with the governance tokens of LPs. More precisely, we encompass the 1-day return, 7-day return, and the daily count of traders engaging with governance tokens. Given that governance tokens bear resemblance to corporate stocks, these three control variables serve to capture certain factors within LPs that may affect protocol performance but are not readily observable.

Second, we incorporate interest rates of the most actively traded cryptocurrencies in decentralized lending markets, as these rates constitute a fundamental determinant of lending system dynamics. Specifically, we select *Wrapped Ether (WETH)* as the representative asset. Notably, even when users engage in identical deposit/borrowing activities for the same cryptocurrency within a single LP, interest rates may vary across users due to protocol-specific rate models (see (Gudgeon et al. (2020b) for detailed mechanism). Consequently, we estimate implied interest rates by aggregating daily transaction records across blockchains for WETH and computing weighted averages based on transaction volumes.[8]

Third, to capture systemic shifts in blockchain finance and their spillover effects on decentralized lending markets, we further introduce three Ether (ETH)-related control variables, given ETH's role as the native asset of the Ethereum blockchain. ETH serves as a foundational asset in DeFi ecosystems, influencing collateral values, transaction costs, and overall market sentiment, which can propagate to lending protocols through mechanisms such as liquidation cascades and capital flows. For instance, empirical studies (e.g., Qin et al. (2021a) and Lehar and Parlour (2022)) have shown that ETH price fluctuations directly impact DeFi lending dynamics by amplifying liquidation risks during volatile periods, as evidenced in analyses of automated liquidation mechanisms and their ties to cryptocurrency volatility. Specifically, we calculate (1) 1-day logarithmic returns, (2) 7-day cumulative returns, and (3) 30-day realized volatility of ETH spot prices. These metrics enable us to disentangle the impact of short-term price movements, medium-term trends, and long-term risk perceptions on LP performance.

Additionally, we consider systemic shocks impacting the decentralized lending market between August 13, 2024, and August 13, 2025. While these events did not originate from our studied protocols, they constitute valid external shocks due to their material impacts on the broader DeFi ecosystem. Specifically, we identify two impact mechanisms. First, we focus on economic shocks causing substantial financial losses across multiple platforms that triggered liquidity crises and contagion effects. Second, we explore regulatory shifts, including European Banking Authority (EBA)'s report and U.S. GENIUS Act, that altered market participants' risk assessments. These events were selected based on three criteria: (1) exogeneity—they were unforeseeable market shocks unrelated to our studied protocols (e.g., sudden hacks); (2) systemic impact—they caused measurable market-wide reactions (e.g., the influences of the U.S. GENIUS Act); and (3) channel specificity—they affected the decentralized lending market through distinct economic/regulatory pathways while leaving our core protocols' financial functions and governance structures intact.

[8] In the supplementary material, we provide a more explicit justification for the use of the weighted average rate of WETH as a robust control variable.

Table 7 Events in the decentralized lending market

Date	Event overview
September 4, 2024	Penpie protocol (built on Pendle) exploited via a security vulnerability in liquid restaking tokens (LRTs), leading to $27.8 million in stolen funds. The attacker targeted yield farming and lending mechanics, but quick intervention by Pendle saved an additional $107 million
September 15, 2024	UwU Lend suffered a $20 million flash loan exploit. Hackers manipulated prices via low-liquidity pools to overborrow assets, draining the protocol in a series of orchestrated transactions
December 3, 2024	Phishing attack on Solana's web3.js library compromised npm packages, leading to ~ $130,000 in drained funds from backend bots and wallets. While not a direct protocol hack, it affected DeFi lending apps on Solana by exposing private keys
January 13, 2025	European Banking Authority (EBA) released a joint report on crypto developments, noting DeFi hacks correlated with market size and stolen assets totalling billions in prior years. It called for enhanced smart contract certification and risk monitoring
March 6, 2025	Bybit's Safe (formerly Gnosis Safe) wallet compromised in a targeted attack bypassing 2FA, leading to unauthorized transactions during emissions adjustments. The hack involved malware displaying fake data while signing poisoned transactions
May 5, 2025	Over $180 million lost in DeFi hacks across multiple protocols, including lending exploits via access control and phishing. Gaming/metaverse integrations saw $290 million breaches, indirectly affecting cross-protocol lending
May 22, 2025	CETUS (SUI's LP provider) hacked, draining liquidity pools and crashing on-chain prices for SUI-based assets. This disrupted DEX-integrated lending on Sui
June 9, 2025	Alex Protocol (Bitcoin DeFi on Stacks) exploited for $8.3 million via a bridge vulnerability, following a prior $4.3 million incident
June 17, 2025	Meta Pool smart contracts exploited, draining 9,700 mpETH worth $27 million. Attackers targeted lending and staking mechanics
June 26, 2025	Resupply stablecoin protocol drained of $9.5 million in a smart contract exploit, exploiting overlooked security in lending code
August 8, 2025	CrediX lending protocol exploited for $4.5 million; team vanished, website offline, suspected exit scam. No reimbursements despite promises
August 6–13, 2025	U.S. GENIUS Act approved, providing regulatory clarity for crypto, including DeFi. This boosted policy-driven market growth

This table introduces several events that have significant impacts on the decentralized lending market from 2024–08-13 to 2025–08-13. More details about the events' influences are presented in the appendix

Table 7 introduces these events, and more details about the events' influences are presented in the appendix. In our analysis, we employ a symmetric ± 2 trading day window around each identified systemic event (2024-08-13 to 2025-08-13) to capture immediate market reactions while minimizing noise in decentralized lending markets characterized by rapid liquidity adjustments. By incorporating event-window dummy variables into the regression frameworks, we rigorously test whether our baseline findings reflect protocol-specific dynamics rather than broader market confounding factors.[9]

Following Saengchote (2023), the independent variables with a one-day lag help to account for potential serial correlation in the data.[10] Standard errors are estimated using the Driscoll and Kraay (1998) procedure. Furthermore, we transform the *net deposit* and *net borrow* variables by taking their first log-differences to ensure stationarity. Notably, the daily panel consists of all protocol–chain units active during the sample period.

[9] This approach is not a conventional event study: we do not estimate abnormal returns, do not define a separate estimation window, and do not attempt to identify a treated versus control group. Instead, following Saengchote (2023), we use these short event windows purely as exogenous shock controls to prevent market-wide volatility from biasing the estimated relationship between mega-whales and protocol outcomes.

[10] The Durbin-Wu-Hausman (DWH) tests for Tables 8, 9, 10 indicate no statistically significant endogeneity issues in the estimation results.

Table 8 Market size dynamics

	(1) ΔIn TVL	(2) ΔIn Total borrow	(3) ΔIn Total deposit	(4) ΔIn Total volume	(5) ΔIn TVL	(6) ΔIn Total borrow	(7) ΔIn Total deposit	(8) ΔIn Total volume
Primary explanatory variables $W_{p,c,t-1}$								
ΔNet deposit	0.00 (1.02)	0.02 (1.30)	−0.10*** (−4.23)	−0.01 (−0.55)	0.00 (1.04)	0.03 (1.39)	−0.09*** (−4.14)	−0.01 (−0.49)
ΔNet borrow	0.00 (0.13)	−0.12*** (−3.09)	−0.04 (−1.37)	−0.05* (−1.91)	0.00 (0.16)	−0.12*** (−3.10)	−0.04 (−1.45)	−0.05* (−1.89)
Txn count	0.00 (0.27)	0.00 (0.07)	0.00 (−0.20)	0.00 (−0.59)	0.00 (0.14)	0.00 (0.05)	0.00 (−0.18)	0.00 (−0.62)
Token count	−0.01*** (−4.15)	−0.02 (−0.40)	−0.02 (−0.52)	−0.04 (−1.13)	−0.01*** (−4.27)	−0.01 (−0.32)	−0.02 (−0.43)	−0.04 (−1.09)
Platform switch ratio	−0.01** (−2.03)	0.23 (1.25)	0.53* (1.92)	0.07 (0.39)	−0.02** (−2.21)	0.24 (1.31)	0.56** (2.00)	0.07 (0.37)
Large trade ratio	0.00 (−0.56)	−0.01 (−0.04)	−0.24 (−1.46)	−0.03 (−0.22)	0.00 (−0.35)	0.01 (0.06)	−0.23 (−1.37)	−0.01 (−0.10)
Borrow utilization	0.00*** (−3.47)	0.00 (−0.71)	0.00 (1.18)	0.00 (−0.75)	0.00*** (−3.43)	0.00 (−0.68)	0.00 (1.18)	0.00 (−0.70)
Control variables X_{t-1}								
Control variables: governance tokens	√	√	√	√	√	√	√	√
Control variables: interest rates	√	√	√	√	√	√	√	√
Control variables: ETH volatility	√	√	√	√	√	√	√	√
Event dummies	×	×	×	×	√	√	√	√
Protocol FE	√	√	√	√	√	√	√	√
N	901	908	908	908	901	908	908	908
Adj R^2	0.34	0.07	0.10	0.03	0.35	0.08	0.11	0.04

This table contains estimates of Eq. (1). Column titles indicate the dependent variable. The dependent variables are strictly scaled as daily log-differences. ΔNet deposit and Δnet borrow are calculated as the log-differences of their respective daily volumes, representing the approximate daily growth rates. Standard errors are estimated using the Driscoll and Kraay (1998) procedure with a one-day lag to account for both cross-sectional and temporal dependence. We exclude year-quarter fixed effects to avoid potential multicollinearity, as market-wide time trends and shocks are effectively captured by the high-frequency Ethereum market variables. t-statistics are in parentheses. $***$, $**$, and $*$ indicate statistical significance at the 1%, 5%, and 10% levels, respectively

Several protocol-chain units exhibit intermittent activity or zero-volume days, which are excluded to ensure consistency in the construction of liquidity and risk ratios. After filtering inactive days and applying lag structure and event-window controls, the final panel contains approximately 1,000 observations, depending on the specification. This is consistent with the expected number of chain-level observations given the sample period and the number of active protocol–chain deployments. The regression models are presented in Eq. (1)[11]:

[11] Equation (1) estimates the short-run associations between mega-whale activity and protocol outcomes. The one-day lag structure accounts for temporal ordering, which helps characterize the lead-lag relationship between the variables. In the supplementary material, we provide Granger-causality checks that are consistent with the hypothesized temporal lead-lag dynamics and potential feedback effects.

Table 9 Structural risk indicators

	(1) ΔDeposit share of volume	(2) ΔDeposit to borrow ratio	(3) ΔDeposit share of volume	(4) ΔDeposit to borrow ratio
Primary explanatory variables $W_{p,c,t-1}$				
ΔNet deposit	−0.01*** (−2.48)	-0.36*** (-3.04)	-0.01** (-2.45)	-0.36*** (-3.05)
ΔNet borrow	0.01*** (2.76)	0.91** (2.22)	0.01*** (2.74)	0.92** (2.23)
Txn count	0.00 (−0.37)	-0.02 (-0.76)	0.00 (-0.38)	-0.02 (-0.77)
Token count	0.00 (−0.93)	-0.04 (-0.07)	0.00 (-0.97)	-0.08 (-0.14)
Platform switch ratio	0.01 (0.16)	-2.26 (-1.62)	0.01 (0.14)	-2.49* (-1.75)
Large trade ratio	−0.01 (−0.77)	-0.07 (-0.05)	-0.01 (-0.75)	-0.05 (-0.04)
Borrow utilization	0.00** (2.16)	0.00 (-0.95)	0.00** (2.16)	0.00 (-0.94)
Control variables X_{t-1}				
Control variables: governance tokens	√	√	√	√
Control variable: interest rate	√	√	√	√
Control variables: ETH volatility	√	√	√	√
Event dummies	×	×	√	√
Protocol FE	√	√	√	√
N	908	908	908	908
Adj $\boldsymbol{R}^2$	0.06	0.03	0.06	0.04

This table contains estimates of Eq. (2). Column titles indicate the dependent variable. The dependent variables are strictly scaled as first differences. ΔNet deposit and Δnet borrow are calculated as the log-differences of their respective daily volumes, representing the approximate daily growth rates. Standard errors are estimated using the Driscoll and Kraay (1998) procedure with a one-day lag to account for both cross-sectional and temporal dependence. We exclude year-quarter fixed effects to avoid potential multicollinearity, as market-wide time trends and shocks are effectively captured by the high-frequency Ethereum market variables. t-statistics are in parentheses.***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

$$Y_{p,c,t} = \alpha + \beta' W_{p,c,t-1} + \gamma' X_{t-1} + \theta' E_t + \delta_p + \varepsilon_{p,c,t} \quad (1)$$

where:

- $Y_{p,c,t}$ denotes the dependent variable(s) of interest related to market size dynamics: *Δln TVL*, *Δln Total borrow*, *Δln Total deposit*, and *Δln Total volume.*
- $W_{p,c,t-1}$ is the vector of mega-whale variables lagged by one day:

$$\begin{aligned} W_{p,c,t-1} = &(\Delta\text{netdeposit}, \Delta\text{netborrow}, \text{txncount}, \text{tokencount}, \\ &\text{platformswitchratio}, \text{largetraderatio}, \text{borrow utilization})_{t-1} \end{aligned}$$

- X_{t-1} is the vector of control variables, including returns and number of traders of governance token, implied interest rates of Wrapped Ether (WETH), and ETH return levels and volatility.
- E_t is the vector of event-window dummy variables indicating the ±2-day windows around major DeFi shocks listed in Table 7.

Table 10 User activity

	(1) ΔIn Unique depositors	(2) ΔIn Unique borrowers	(3) ΔIn Total unique users	(4) ΔIn Txn count protocol	(5) ΔIn Unique depositors	(6) ΔIn Unique borrowers	(7) ΔIn Total unique users	(8) ΔIn Txn count protocol
Primary explanatory variables $W_{p,c,t-1}$								
ΔNet deposit	0.01* (1.82)	0.00 (0.37)	0.00 (0.81)	0.00 (0.90)	0.01* (1.79)	0.00 (0.42)	0.00 (0.77)	0.00 (0.93)
ΔNet borrow	0.00 (− 0.18)	0.00 (− 0.39)	0.00 (− 0.77)	− 0.01 (− 0.99)	0.00 (− 0.24)	0.00 (− 0.45)	0.00 (− 0.77)	− 0.01 (− 1.02)
Txn count	0.00 (− 0.42)	0.00 (− 0.92)	0.00 (− 0.74)	0.00 (− 1.29)	0.00 (− 0.48)	0.00 (− 0.96)	0.00 (− 0.86)	0.00 (− 1.33)
Token count	0.01 (0.77)	0.00 (0.22)	0.00 (− 0.05)	0.01 (0.81)	0.01 (0.84)	0.00 (0.29)	0.00 (0.00)	0.01 (0.89)
Platform switch ratio	0.20** (2.12)	0.07 (1.16)	0.07 (1.05)	0.12* (1.76)	0.21** (2.21)	0.04 (1.23)	0.07 (1.08)	0.12* (1.85)
Large trade ratio	− 0.01 (− 0.13)	0.00 (0.08)	0.00 (0.06)	0.02 (0.56)	− 0.01 (− 0.14)	0.01 (0.16)	0.01 (0.10)	0.03 (0.60)
Borrow utilization	0.00 (− 0.78)	0.00 (− 0.77)	0.00 (− 0.77)	0.00 (− 0.97)	0.00 (− 0.76)	0.00 (− 0.72)	0.00 (− 0.71)	0.00 (− 0.97)
Control variables X_{t-1}								
Control variables: governance tokens	√	√	√	√	√	√	√	√
Control variable: interest rate	√	√	√	√	√	√	√	√
Control variables: ETH volatility	√	√	√	√	√	√	√	√
Event dummies	×	×	×	×	√	√	√	√
Protocol FE	√	√	√	√	√	√	√	√
N	908	908	908	908	908	908	908	908
Adj R^2	0.03	0.02	0.02	0.03	0.05	0.03	0.04	0.03

This table contains estimates of Eq. (3). Column titles indicate the dependent variable. The dependent variables are strictly scaled as daily log-differences. ΔNet deposit and Δnet borrow are calculated as the log-differences of their respective daily volumes, representing the approximate daily growth rates. Standard errors are estimated using the Driscoll and Kraay (1998) procedure with a one-day lag to account for both cross-sectional and temporal dependence. We exclude year-quarter fixed effects to avoid potential multicollinearity, as market-wide time trends and shocks are effectively captured by the high-frequency Ethereum market variables. t-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

- δ_p are protocol fixed effects.[12]
- $\varepsilon_{p,c,t}$ is the error term.

[12] We intentionally exclude year-quarter fixed effects from our baseline specification to mitigate potential multicollinearity with our high-frequency, time-varying market controls (e.g., daily WETH implied rates, ETH volatility) and specific DeFi shock event dummies. This specification choice is strongly supported by empirical diagnostics. Specifically, incorporating year-quarter fixed effects leads to a noticeable reduction in the adjusted R^2, suggesting that these low-frequency dummies over penalize the model's degrees of freedom without contributing meaningful incremental explanatory power. Furthermore, as evidenced by the specification sensitivity analysis in Table 8, the exclusion of these fixed effects prevents artificial variance inflation, thereby stabilizing our core coefficient estimates and ensuring the robustness of our empirical inferences.

Table 8 presents the empirical findings from Eq. (1), which investigate the relationship between mega-whale activities and market size dynamics in LPs. The dependent variables include daily log-differences in TVL (*Δln TVL*), Total Borrow (*Δln Total borrow*), Total Deposit (*Δln Total deposit*), and Total Volume (*Δln Total volume*).

The token count (daily traded token species by mega-whales) exhibits a statistically significant negative relationship with *Δln TVL*, and the result is significant at the 1% level. This indicates that each additional traded token species correlates with a 1% decline in TVL. Economically, this suggests that fragmented trading across multiple assets may reduce protocol-level value accumulation. Such diversification could reflect arbitrage strategies that offset positions across tokens, limiting sustained impacts on individual protocols.

The *platform switch ratio* implies the cross-protocol arbitrage intensity of mega-whales, and it demonstrates a statistically significant negative association with *Δln TVL*. This implies that frequent protocol switching by mega-whales constrains TVL growth in any single protocol. While cross-protocol activity may transiently boost transactional volumes, it does not meaningfully expand protocol-specific liquidity pools, aligning with the notion of arbitrage-driven liquidity fragmentation. This empirical finding directly aligns with the behavioral profile of Cluster 1 (Arbitrageur/Cross-protocol Reallocator) identified in Sect. "Different types of whale users in decentralized lending". These users prioritize opportunistic reallocation across platforms to optimize yields, and their high switching frequency prevents the formation of stable, long-term liquidity buffers within a single protocol.

ΔNet borrow is strongly and negatively correlated with both *Δln total borrow* and *Δln total volume*. This evidence suggests that higher leverage ratios, when driven by mega-whale activity, are associated with diminished growth in borrowing and overall protocol volume. Such leveraged strategies likely exert a negative externality on the protocol by crowding out or reducing broader financial activities. In essence, while individual mega-whales may maximize leverage, their behavior may ultimately attenuate protocol-level value accumulation and long-term scalability.

The *ΔNet deposit* shows a strong negative correlation with the *Δln total deposit*. Specifically, a 1 percentage point increase in the growth rate of net deposits is associated with a 0.10 percentage point decrease in the growth rate of total deposits, as evidenced by the statistically significant coefficient of -0.10 in Column 3, Table 8. While liquidity provision is vital for the viability of DeFi ecosystems, excessive concentration can introduce significant risks. The growth of total deposits tends to decelerate when mega-whales dominate liquidity provision.

The results remain consistent after including DeFi event dummies. This supports the robustness of our original findings by demonstrating their persistence even when controlling for major DeFi ecosystem disruptions. Furthermore, they strengthen the empirical associations by showing that mega-whale activities precede and are strongly associated with observed outcomes, even when controlling for market-wide external shocks. This methodological approach, with its short event window and explicit control for systemic events, enhances internal validity by helping mitigate alternative explanations based on market-wide phenomena while preserving the causal logic linking decentralized LPs' design features to empirical outcomes.

Collectively, the regression results partially support Research Hypothesis 1.[13] Mega-whale directional bets (*Δnet deposit* and *Δnet borrow*), together with their transaction frequency and potential arbitrage activities, emerge as significant lead indicators of liquidity expansion. These findings underscore the nuanced role of whale behavior: mega-whales' activities may negatively affect LPs, but speculative transactions (indicated by the *platform switch ratio*) can contribute positively to the growth of total deposits in LPs.

Structural risk indicators

In this section, we investigate how mega-whales influence structural risk in LPs.[14] Consistent with Sect. "Market size dynamics", we estimate standard errors using the Driscoll and Kraay (1998) procedure with a one-day lag to account for potential serial correlation in the data, and the independent variables and other settings of regression models are the same. The analysis is conducted at the protocol-chain level, an approach justified by the fact that DeFi protocols can be deployed across multiple blockchain networks. The regression models are presented in Eq. (2)[15]:

$$Y_{p,c,t} = \alpha + \beta' W_{p,c,t-1} + \gamma' X_{t-1} + \theta' E_t + \delta_p + \varepsilon_{p,c,t} \tag{2}$$

where:

- $Y_{p,c,t}$ denotes the dependent variable(s) of interest related to structural risk indicators: *ΔDeposit share of volume*, *ΔDeposit to borrow ratio*.
- $W_{p,c,t-1}$ is the vector of mega-whale variables lagged by one day:

$$\begin{aligned} W_{p,c,t-1} = &(\Delta \text{netdeposit}, \Delta \text{netborrow}, \ \text{txncount}, \text{tokencount}, \\ &\text{platformswitch ratio}, \text{large traderatio}, \text{borrow utilization})_{t-1} \end{aligned}$$

- X_{t-1} is the vector of control variables, including returns and number of traders of governance token, implied interest rates of Wrapped Ether (WETH), and ETH return levels and volatility.
- E_t is the vector of event-window dummy variables indicating the ± 2-day windows around major DeFi shocks listed in Table 7.

[13] The adjusted R^2 values in Tables 8, 9, 10 are low, which is typical for high-frequency DeFi data. Daily liquidity flows are driven by many idiosyncratic factors—price shocks, liquidation events, cross-chain arbitrage, gas-fee spikes—that are not fully observable. Our model is not intended to explain total variation, but to identify systematic components associated with mega-whale activity. Low R^2 is therefore consistent with a setting where explanatory variables capture short-run dynamics, not long-term structural determinants of liquidity. For instance, Liu and Tsyvinski (2021) document similarly low R^2 levels in their factor analysis of cryptocurrency returns, emphasizing that while predictive power (as measured by R^2) may be limited, the statistical significance of key risk factors remains robust.

[14] Structural risk in DeFi lending can also be characterized using collateral-based measures such as collateral volatility, collateral diversification, or pool reserves relative to protocol-wide reserves. However, these measures require pool-level collateral composition data at high frequency for each protocol–chain pair, which is not uniformly available in Flipside datasets and is often updated irregularly in on-chain metadata. Moreover, different protocols implement heterogeneous collateral models (e.g., isolated pools, cross-collateral models, stablecoin-specific risk frameworks), making these measures difficult to aggregate in a way that remains comparable across protocols and blockchains. For these reasons, we focus on liquidity-based risk indicators that are directly and consistently observable across all protocols and that map cleanly into solvency and utilization risk—two key channels through which megawhale behavior can influence overall protocol functioning.

[15] As with Eq. (1), these results show short-run associations between mega-whale activity and protocol outcomes. The one-day lag structure ensures temporal ordering, which strengthens the empirical basis for analyzing lead-lag dynamics. In the supplementary material, we provide additional lag specifications and Granger-causality checks that further illustrate these predictive relationships and potential feedback effects.

- δ_p are protocol fixed effects.
- $\varepsilon_{p,c,t}$ is the error term.

Table 9 examines the impact of mega-whale behavior on structural risk in LPs, focusing on *Δdeposit share of volume* (measuring liquidity concentration) and *Δdeposit to borrow ratio* (measuring leverage stability) by estimating Eq. (2).

ΔNet deposit exhibits a statistically significant negative relationship with both structural risk indicators. Specifically, a one-unit increase in *ΔNet deposit* is associated with a 0.01 percentage point decline in *Δdeposit share of volume* (significant at the 1% level) and a 0.36 percentage point reduction in *Δdeposit to borrow ratio* (significant at the 1% level). These findings indicate that elevated whale deposits fragment liquidity concentration and destabilize leverage balances.

ΔNet borrow shows a positive effect on *Δdeposit share of volume*. Specifically, a one-unit increase in *Δnet borrow* is associated with a 0.01 percentage point increase in the deposit share of volume (significant at the 1% level) and a 0.92 percentage point increase in the deposit-to-borrow ratio (significant at the 5% level).[16] This aligns with Hypothesis 2, suggesting a link between borrowing-related arbitrage and the attraction of complementary deposits, which temporarily boosts liquidity concentration. The stabilizing effect of whale borrowing also aligns with the 'strategic inertia' identified in Cluster 2 (high-throughput whales) and Cluster 4 (long-term leverage users), who maintain high borrow shares (0.40 and 0.35, respectively). Unlike the transient deposits of arbitrageurs, the borrowing positions of these archetypes represent more committed, long-term leverage strategies, which help balance the protocol's utilization ratios and mitigate liquidity fragmentation.

Arbitrage-related metrics yield positive but marginally significant results. The *platform switch ratio* negatively impacts the *Δdeposit to borrow ratio*, suggesting that cross-LP arbitrage may dilute the stability of LPs.

In summary, the results offer partial confirmation of Research Hypothesis 2.[17] Mega-whale deposits exert a negative impact on structural risk by fragmenting liquidity and destabilizing leverage balances, whereas their borrowings positively influence deposit share, suggesting that borrowing-driven activities temporarily bolster liquidity concentration. Arbitrage strategies further shape risk outcomes. While the model's low adjusted R^2 values—likely attributable to noise in high-frequency risk metrics—limit explanatory power, these findings underscore the dual role of mega-whales: as stabilizers through borrowing-fuelled arbitrage and as destabilizers via deposit-driven fragmentation or leveraged risk-taking.

[16] While the estimated daily effect of 0.01 percentage points appears small, its cumulative impact is economically meaningful. For instance, if sustained, this corresponds to approximately a 0.3% decline over 30 days and nearly 1.8% over six months. Given the scale of TVL in leading protocols (often billions of USD), such cumulative effects translate into economically significant capital reallocations.

[17] The rationale for the observed R^2 levels is discussed in detail in Sect. "Market size dynamics" (see Footnote 13).

User activity

In this section, we investigate how mega-whales influence user activity in LPs. Consistent with the previous sections, we include independent variables with a one-day lag to account for potential serial correlation in the data, and the independent variables and other settings of regression models are the same. The analysis is conducted at the protocol-chain level, an approach justified by the fact that DeFi protocols can be deployed across multiple blockchain networks. The regression models are presented in Eq. (3):

$$Y_{p,c,t} = \alpha + \beta' W_{p,c,t-1} + \gamma' X_{t-1} + \theta' E_t + \delta_p + \varepsilon_{p,c,t} \tag{3}$$

where:

- $Y_{p,c,t}$ denotes the dependent variable(s) of interest related to user activity: *Δln Unique depositors, Δln Unique borrowers, Δln Total unique users, Δln Txn count protocol.*
- $W_{p,c,t-1}$ is the vector of mega-whale variables lagged by one day:

$$\begin{aligned} W_{p,c,t-1} = & (\Delta \text{netdeposit}, \Delta \text{netborrow}, \text{txncount}, \text{tokencount}, \\ & \text{platformswitchratio}, \text{largetraderatio}, \text{borrowutilization})_{t-1} \end{aligned}$$

- X_{t-1} is the vector of control variables, including returns and number of traders of governance token, implied interest rates of Wrapped Ether (WETH), and ETH return levels and volatility.
- E_t is the vector of event-window dummy variables indicating the ± 2-day windows around major DeFi shocks listed in Table 7.
- δ_p are protocol fixed effects.
- $\varepsilon_{p,c,t}$ is the error term.

Table 10 examines the impact of mega-whale behavior on user activity in LPs, focusing on four dependent variables: *Δln Unique depositors*, *Δln Unique borrowers*, *Δln Total unique users*, and *Δln Txn count protocol*.

ΔNet deposit exhibits a positive relationship with user activity in LPs. Specifically, a 1% increase in net deposits is associated with a 0.01% increase in unique depositors. This suggests that elevated whale deposits may attract retail users. The positive correlation between whale deposits and retail participation may be driven by the presence of Cluster 3 (passive-treasury style whales). Although these users have low transaction intensity, their large-scale, stable liquidity provision signals protocol safety to the broader market. This finding reinforces the 'rational herding' hypothesis, where retail users view the entry of such conservative mega-whales as a proxy for institutional-grade security.

The *platform switch ratio* also shows positive effects on user activities. A 1 percentage point increase in the *platform switch ratio* is associated with a 0.20% increase in the number of *unique depositors* and a 0.12% increase in the *txn count protocol*. These findings suggest that the cross-LP activity of mega-whales also benefits the underlying LPs. This is potentially because depositors realize that providing liquidity attracts whales, allowing them to earn higher profits; consequently, transaction counts on these platforms grow more rapidly.

To summarize, the results partially support Research Hypothesis 3.[18] Mega-whale deposits are positively associated with user activity, likely by attracting retail participants, while cross-LP activities positively influence both depositor counts and transaction counts. These findings highlight the interplay between whale activity and user participation, with whales acting as attractors for retail involvement.

The empirical evidence reveals a dual-edged influence of mega-whale activities on protocol liquidity and user dynamics; however, such influence does not inherently equate to a direct estimate of systemic risk or governance concentration. Consequently, we interpret these findings as preliminary signals of potential vulnerability rather than conclusive proof of systemic instability, acknowledging that whale participation may yield both benefits and risks to the underlying platforms. Any references to broader governance or systemic implications should be understood as contextual interpretations grounded in prior literature rather than causal findings established by the regressions.[19]

Conclusion

This study empirically analyses the influence of mega-whales—users with transaction volumes exceeding $100 million—on LPs such as *Aave v2/v3*, *Compound v2/v3*, *Morpho Blue*, *Flux*, and *SparkLend*, using high-frequency transaction-level data from Flipside Crypto spanning August 2024 to August 2025. Identifying 550 such whales, we find that their activities exert a dual impact: Empirical results reveal that the growth in mega-whales' net deposits and borrowing positions is negatively correlated with overall protocol expansion. Furthermore, arbitrage-driven behaviors, such as token diversification and platform switching, exert a fragmenting effect on TVL growth, reflecting statistically significant conditional decreases in TVL growth. Conversely, platform switching demonstrates a significant positive correlation with the growth of total deposits. Regarding structural stability and user dynamics, borrowing activities serve as temporary stabilizers for leverage ratios, while mega-whale deposits are found to catalyze retail participation, as evidenced by positive effects on the growth of unique depositors. Event studies around DeFi shocks (e.g., Penpie exploit) and regulatory shifts reveal a reflexive relationship where whale behaviors and protocol responses are deeply intertwined, suggesting a feedback loop rather than a simple exogenous shock. These results suggest patterns that are consistent with concerns raised in prior work about economic concentration in DeFi. While our empirical focus is on liquidity and user activity, such patterns may have implications for broader decentralization debates (Sun et al. 2024).

Theoretically, our results illuminate key principles and regularities in decentralized markets, extending classical financial intermediation theories to the blockchain context. First, they reveal a concentration-liquidity trade-off in LPs, where mega-whales act as both liquidity enhancers and risk amplifiers, echoing Diamond and Dybvig's (1983) bank run models but adapted to algorithmic, permissionless

[18] Consistent with the previous analysis, the R^2 remains low due to the high-frequency nature of the dataset.

[19] While our analysis primarily focuses on the 'whale-to-protocol' channel, the Granger causality tests in the supplementary material (Table OA.7) indicate a bidirectional feedback loop. This suggests that while whales act as catalysts, they also respond dynamically to protocol conditions.

environments. For example, speculative fragmentation—evident in negative coefficients for token count (− 0.01 on *Δln TVL*) and *platform switch ratio* (− 0.01 on *Δln TVL*)—highlights how arbitrage incentives erode systemic stability, as modeled in Milionis et al. (2025). Second, the results indicate the stabilizing role of borrowing on structural risks (e.g., positive effects of growth of net borrowing positions on the growth deposit share of volume), and the growth of mega-whale deposits show positive influences on user activity (e.g., 0.01 elasticity on the growth unique depositors). The findings demonstrate a feedback loop between whale leverage and protocol resilience, revealing regularities in herd behavior and risk contagion (Sockin and Xiong 2023). Overall, these insights affirm that while DeFi achieves technical decentralization through distributed governance, its purported broad-based decentralization is often illusory due to economic concentration, with mega-whales exerting outsized influence comparable to institutional investors in centralized systems (Chernoff and Jagtiani 2025), thereby contributing to the literature on DeFi governance and systemic fragility (Gudgeon et al. 2020b; Chiu et al. 2023; Sun et al. 2024).

In practical terms, our findings offer actionable implications for stakeholders within the evolving FinTech landscape. For protocol designers, the evidence that mega-whale deposits positively stimulate retail participation suggests that while whales attract users, their concentrated positions simultaneously fragment TVL. This dual impact underscores the need for sophisticated risk management frameworks, such as dynamic utilization caps, real-time whale monitoring via on-chain analytics, and incentive mechanisms designed to diversify liquidity sources and mitigate liquidation cascades. Regulators and policymakers can leverage these results to inform balanced approaches to DeFi oversight—for instance, incorporating whale activity monitoring into frameworks such as the U.S. GENIUS Act. Such oversight is critical to preventing the excessive concentration that exacerbates vulnerabilities, as seen in the Penpie exploit (September 2024). For retail users, understanding the reflexive relationship between whale activity and protocol resilience is essential, while they may follow whale-induced liquidity, they must also hedge against the volatility and risk contagion inherent in concentrated markets. Ultimately, by quantifying the impacts of the 550 identified mega-whales, this study aids in fostering more resilient DeFi ecosystems and reducing potential financial losses attributed to liquidity and concentration risks.

Compared to prior research, our work builds upon and extends existing studies in several ways. Unlike Gudgeon et al. (2020b) and Bartoletti et al. (2021), who provide theoretical models of DeFi lending risks without large-scale empirical evidence on user concentration, we offer granular, transaction-level analysis across multiple protocols and blockchains, filling a critical gap in behavioral impacts (Saengchote 2023). We also bridge traditional banking literature—e.g., depositor discipline (Berger and Turk-Ariss 2014; Santos and Winton 2019)—to DeFi by demonstrating parallels in liquidity risks and market power, but with corrections for blockchain-specific features such as anonymity and flash loans, which amplify whale arbitrage (Wu et al. 2025; Capponi and Jia 2025). While Chernoff and Jagtiani (2025) focus on general crypto whale behaviors, our study supplements this by exploring the reciprocal interactions between mega-whales and specific protocols, revealing nuances such as the positive borrowing-user activity link absent in broader market analyses. These developments refine the decentralization illusion narrative (Sun et al. 2024), providing empirical substantiation that challenges overly optimistic views of DeFi's inclusivity (Gogel et al. 2021; Ali 2024).

Although we discuss systemic and governance-related implications, our empirical analysis is limited to liquidity dynamics, user activity, and structural ratios at the pool–chain level. We therefore interpret broader implications as suggestive rather than conclusive, and future work incorporating direct measures of liquidation cascades, governance voting power, or cross-protocol contagion could more fully establish systemic risk channels.

Finally, our research has limitations that point to avenues for future inquiry. Anonymity in blockchain data precludes the direct identification of mega-whales, limiting deeper insights into their strategies (e.g., institutional vs. individual actors), although methods such as Frijns et al. (2022) could be adapted for geolocation or network analysis. Additionally, our metrics do not fully capture social coordination among whales or off-chain influences, potentially underestimating illiquidity attack risks (Kulkarni et al. 2022). Future studies might integrate machine learning for predictive modeling of whale-driven cascades or explore cross-chain effects post-regulatory harmonization (e.g., under the GENIUS Act). Moreover, extending the analysis to emerging LPs on non-EVM chains (e.g., Solana) or incorporating real-time oracle data could enhance generalizability. In sum, while DeFi promises financial inclusion, our findings urge a cautious approach: mitigating whale concentration is essential to realizing its transformative potential without replicating traditional finance inequities.

Appendix

This appendix includes further technical information complementing the main text of this study. See Tables 11 and 12.

Table 11 Abbreviations

Abbreviation	Full Form
LP	Lending Protocol

Table 12 Event impacts

Date	Impact on LPs
September 4, 2024 (Penpie protocol)	Eroded trust in yield-enhanced lending; highlighted risks in composable DeFi layers. Protocols like Pendle saw temporary TVL drops, prompting audits and slower adoption of LRT-integrated lending. Overall, it reinforced the need for better oracle and contract security in lending pools
September 15, 2024 (UwU Lend)	Major hit to lending confidence; UwU's TVL plummeted, affecting users' collateral. It accelerated industry-wide flash loan mitigations, like rate limits and multi-oracle checks, but contributed to broader 2024 DeFi losses exceeding $313 million in August alone
December 3, 2024 (Solana's web3.js library)	Indirect impact on Solana-based lending (e.g., via bots managing loans); devs updated packages urgently, but it underscored supply-chain risks. No major protocol TVL loss, but heightened scrutiny on client-side security for lending interfaces
January 13, 2025 (European Banking Authority)	Regulatory pressure increased; EU protocols faced compliance reviews, slowing lending innovations but improving long-term security. Contributed to a 40% drop in DeFi losses for 2024 overall, per Hacken reports
March 6, 2025 (Bybit's Safe)	Exposed frontend vulnerabilities in multisig lending ops; affected institutional lending flows. Pushed for decentralized frontends and better UX in transaction verification, impacting protocols relying on Safe for treasury/lending management
May 5, 2025 (Over $180 million lost in DeFi hacks)	Systemic risk spike; lending TVL dipped as users withdrew collateral. Highlighted need for robust access controls, leading to industry-wide bounties and audits for lending platforms
May 22, 2025 (CETUS)	Severe liquidity crunch in Sui DeFi lending; protocols paused operations, causing cascading liquidations. It amplified calls for chain-specific security, reducing Sui lending adoption temporarily
June 9, 2025 (Alex Protocol)	Damaged Bitcoin DeFi credibility; TVL in Stacks lending protocols fell sharply. Team's reimbursement pledge mitigated some loss, but it deterred capital inflows to emerging BTC lending ecosystems
June 17, 2025 (Meta Pool)	Major blow to ETH restaking/lending; protocols like Meta Pool halted services, leading to user panic withdrawals. Contributed to mid-2025's hack wave, pushing for recursive security testing
June 26, 2025 (Resupply stablecoin protocol)	Undermined stablecoin lending trust; highlighted TVL-security mismatches. Resulted in broader audits, but added to 2025's $1.1 billion total DeFi losses
August 8, 2025 (CrediX)	Devastating for user trust; CrediX's TVL zeroed out, amplifying scam fears in smaller lending protocols. Led to calls for better team verification and insurance mandates
August 6–13, 2025 (U.S. GENIUS Act)	Positive shock; increased institutional participation in DeFi lending, with TVL rising post-approval. However, it raised compliance costs for protocols, potentially excluding smaller players

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40854-026-00956-3.

Additional file1 (DOCX 35 KB)

Acknowledgements
None.

Author contributions
Xiaotong Sun: Conceptualization, Methodology, Formal Analysis, Writing – Original Draft. Charalampos Stasinakis: Formal Analysis, Supervision, Writing – Review & Editing. Georgios Sermpinis: Supervision, Writing – Review & Editing.

Funding
Open Access funding enabled and organized by CAUL and its Member Institutions. This research received no specific grant from any funding agency.

Data availability
The datasets generated during the study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Received: 9 September 2025 Revised: 27 June 2026 Accepted: 29 June 2026
Published online: 22 July 2026

References
Ali A (2024) Decentralized finance (DEFI) and its impact on traditional banking systems: opportunities, challenges, and future directions. SSRN Electron J. https://doi.org/10.2139/ssrn.4942313
An J, Rau R (2019) Finance, technology and disruption. Eur J Finance 27:334–345. https://doi.org/10.1080/1351847x.2019.1703024
Aramonte S, Doerr S, Huang W, Schrimpf A (2022) DeFi Lending: Intermediation without Information? (No. 57). BIS Working Papers
Auer R, Claessens S (2018) Regulating cryptocurrencies: assessing market reactions. BIS Quarterly Review
Azar PD, Casillas A, Farboodi M (2024) Information and market power in DeFi intermediation (No. w32949). National Bureau of Economic Research
Bartoletti M, Chiang JH-Y, Lafuente AL (2021) SOK: lending pools in decentralized finance. In: Lecture notes in computer science. pp 553–578
Berger AN, Turk-Ariss R (2014) Do depositors discipline banks and did government actions during the recent crisis reduce this discipline? An international perspective. J Financ Serv Res 48:103–126. https://doi.org/10.1007/s10693-014-0205-7
Capponi A, Jia R (2025) Liquidity provision on blockchain-based decentralized exchanges. Rev Financ Stud. https://doi.org/10.1093/rfs/hhaf046
Capponi A, Iyengar G, Sethuraman J (2023) Decentralized finance: protocols, risks, and governance. Found Trends® Priv Secur 5:144–188. https://doi.org/10.1561/3300000036
Castro-Iragorri C, Ramirez J, Velez S (2021) Financial intermediation and risk in decentralized lending protocols. SSRN Electron J. https://doi.org/10.2139/ssrn.3893278
Chernoff A, Jagtiani J (2025) Beneath the crypto currents: the hidden effect of crypto "whales." J Credit Risk. https://doi.org/10.21314/jcr.2025.007
Chiu J, Ozdenoren E, Yuan K, Zhang S (2023) On the fragility of DeFi lending. SSRN Electron J. https://doi.org/10.2139/ssrn.4328481
Cornelli G, Gambacorta L, Garratt R, Reghezza A (2025) Why DeFi lending? Evidence from Aave V2. J Financ Intermed. https://doi.org/10.1016/j.jfi.2025.101166
Diamond DW, Dybvig PH (1983) Bank runs, deposit insurance, and liquidity. J Polit Econ 91:401–419. https://doi.org/10.1086/261155
Driscoll JC, Kraay AC (1998) Consistent covariance matrix estimation with spatially dependent panel data. Rev Econ Stat 80:549–560
Feinstein BD, Werbach K (2021) The impact of cryptocurrency regulation on trading markets. J Financ Regul 7:48–99. https://doi.org/10.1093/jfr/fjab003
Frijns BPM, Foley S, Garel A, Roh T-Y (2022) Who buys Bitcoin? The cultural determinants of Bitcoin activity. Int Rev Financ Anal 84:102385. https://doi.org/10.1016/j.irfa.2022.102385
Gogel D, Deshmukh S, Geest A, Resas D, Sillaber C, Werbach K (2021) DeFi Beyond the hype: the emerging world of decentralized finance. In: Collaborative publication with Wharton Blockchain and Digital Asset Project and World Economic Forum. Wharton, Philadelphia. https://wifpr.wharton.upenn.edu/wp-content/uploads/2021/05/DeFi-Beyond-the-Hype.pdf
Gudgeon L, Perez D, Harz D, et al (2020a) The decentralized financial crisis. Proceedings of 2020 Crypto Valley conference on blockchain technology (CVCBT), pp. 1–15. https://doi.org/10.1109/cvcbt50464.2020.00005
Gudgeon L, Werner S, Perez D, Knottenbelt WJ (2020b) DeFi protocols for loanable funds. Proceedings of the 2nd ACM conference on advances in financial technologies. https://doi.org/10.1145/3419614.3423254
Han J, Lee J, Li T (2025) A review of DAO governance: recent literature and emerging trends. J Corp Financ 91:102734
Harvey CR, Ramachandran A, Santoro J (2021) DEFI and the future of finance. John Wiley & Sons
Heimbach H, Huang W (2024) DeFi leverage (No. 1184). BIS Working Papers
Kaplan B, Benli VF, Alp EA (2023) Decentralized finance and new lending protocols. PressAcademia Procedia 16(1):192–195
Kulkarni K, Diamandis T, Chitra T (2022) Towards a theory of maximal extractable value I: constant function market makers. In: arXiv.org. https://arxiv.org/abs/2207.11835
Lehar A, Parlour CA (2022) Systemic fragility in decentralised markets. https://www.bis.org/publ/work1062.htm
Liu Y, Tsyvinski A (2021) Risks and returns of cryptocurrency. Rev Financ Stud 34:2689–2727. https://doi.org/10.1093/rfs/hhaa113

Liu J, Li X, Wang S (2020) What have we learnt from 10 years of Fintech research? A scientometric analysis. Technol Forecast Soc Change 155:120022. https://doi.org/10.1016/j.techfore.2020.120022
McLaughlin R, Kruegel C, Vigna G (2023) A large scale study of the ethereum arbitrage ecosystem. In: 32nd USENIX security symposium (USENIX Security 23). pp 3295–3312
Milionis J, Moallemi CC, Roughgarden T (2025) Automated market making and arbitrage profits in the presence of fees. In: Kiayias A, Meiklejohn S (eds) Lecture notes in computer science. Springer, Cham, pp 159–171
Papangelou S, Christodoulou K, Michoulis G (2023) Exploring decentralized governance: a framework applied to compound finance. In: the international conference on mathematical research for blockchain economy. Springer Nature Switzerland, Cham, pp 152–168
Pritsker M (2002) Large investors and liquidity: a review of the literature. Risk measurement and systematic risk—proceedings of the third joint central bank research conference. Bank for International Settlements, Basel, pp 126–139
Qin K, Zhou L, Afonin Y, et al (2021a) CeFi vs. DeFi -- Comparing centralized to decentralized finance. In: arXiv.org. https://arxiv.org/abs/2106.08157
Qin K, Zhou L, Gamito P, Jovanovic P, Gervais A (2021b) An empirical study of DeFi liquidations: incentives, risks, and instabilities. Proceedings of the 21st ACM internet measurement conference, pp 336–350
Saengchote K (2023) Decentralized lending and its users: insights from Compound. J Int Financ Mark Inst Money 87:101807. https://doi.org/10.1016/j.intfin.2023.101807
Sai AR, Buckley J, Fitzgerald B, Gear AL (2021) Taxonomy of centralization in public blockchain systems: a systematic literature review. Inf Process Manag 58(4):102584. https://doi.org/10.1016/j.ipm.2021.102584
Santos JaC, Winton A (2019) Bank capital, borrower power, and loan rates. Rev Financ Stud 32:4501–4541. https://doi.org/10.1093/rfs/hhz001
Sockin M, Xiong W (2023) A model of cryptocurrencies. Manag Sci 69:6684–6707. https://doi.org/10.1287/mnsc.2023.4756
Soros G (1987) The alchemy of finance: reading the mind of the market. Simon & Schuster
Sun X, Stasinakis C, Sermpinis G (2024) Decentralization illusion in decentralized finance: evidence from tokenized voting in MakerDAO polls. J Financ Stabil 73:101286. https://doi.org/10.1016/j.jfs.2024.101286
Wang D, Wu S, Lin Z, et al (2021) towards a first step to understand flash loan and its applications in DeFi ecosystem. Proceedings of the ninth international workshop on security in blockchain and cloud computing. https://doi.org/10.1145/3457977.3460301
Weingärtner T, Fasser F, Da Costa PRS, Farkas W (2023) Deciphering DEFI: a comprehensive analysis and visualization of risks in decentralized finance. J Risk Financial Manag 16:454. https://doi.org/10.3390/jrfm16100454
Werner S, Perez D, Gudgeon L, et al (2022) SoK: decentralized finance (DeFi). Proceedings of the 4th ACM conference on advances in financial technologies. https://doi.org/10.1145/3558535.3559780.
Wu Z, Wu J, Zhang H, et al (2025) Hunting in the dark forest: a pre-trained model for on-chain attack transaction detection in Web3. Proceedings of WWW '25: proceedings of the ACM on web conference 2025 4519–4530. https://doi.org/10.1145/3696410.3714928
Zhang J, Liu P (2012) Rational herding in microloan markets: evidence from Prosper.com. Inf Syst Res 23:895–912

Publisher's Note