DeFi Survival Analysis: Insights Into the Emerging Decentralized Financial Ecosystem

AARON M. GREEN, Rensselaer Polytechnic Institute, USA
MICHAEL P. GIANNATTASIO, Rensselaer Polytechnic Institute, USA
JOHN S. ERICKSON, Rensselaer Polytechnic Institute, USA
OSHANI SENEVIRATNE, Rensselaer Polytechnic Institute, USA
KRISTIN P. BENNETT, Rensselaer Polytechnic Institute, USA

We propose a survival analysis approach for discovering and characterizing user behavior and risks for lending protocols in decentralized finance (DeFi). We demonstrate how to gather and prepare DeFi transaction data for survival analysis. We illustrate our approach using transactions in Aave, one of the largest lending protocols. We develop a DeFi survival analysis pipeline that first prepares transaction data for survival analysis through the selection of different index events (or transactions) and associated outcome events. Then we apply survival analysis statistical and visualization methods such as median survival times, Kaplan–Meier survival curves, and Cox hazard regression to gain insights into usage patterns and risks within the protocol. We show how, by varying the index and outcome events as well as covariates, we can use DeFi survival analysis to answer questions like "How does loan size affect the repayment schedule of the loan?"; "How does loan size affect the likelihood that an account gets liquidated?"; "How does user behavior vary between Aave markets?"; "How has user behavior in Aave varied from quarter to quarter?" The proposed DeFi survival analysis can easily be generalized to other DeFi lending protocols. By defining appropriate index and outcome events, DeFi survival analysis can be applied to any cryptocurrency protocol with transactions.

CCS Concepts: • **Mathematics of computing** → **Mathematical analysis**; • **Applied computing** → **Electronic commerce**.

Additional Key Words and Phrases: decentralized finance, aave, survival analysis, financial technologies, fintech, economics, data analytics

ACM Reference Format:
Aaron M. Green, Michael P. Giannattasio, John S. Erickson, Oshani Seneviratne, and Kristin P. Bennett. 2023. DeFi Survival Analysis: Insights Into the Emerging Decentralized Financial Ecosystem. 1, 1 (June 2023), 22 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

The rise in popularity of Bitcoin in the last decade has brought with it the novel study of blockchain technologies in both academic and industrial environments. An accompanying new financial ecosystem has started to emerge, called Decentralized Finance (DeFi). Built using smart contracts (code deployed onto blockchains), the infancy of this novel financial market has seen developers attempting to recreate the functions of traditional, centralized financial services in

Authors' addresses: Aaron M. Green, greena12@rpi.edu, Rensselaer Polytechnic Institute, 110 8th St, Troy, New York, USA, 12180; Michael P. Giannattasio, giannm@rpi.edu, Rensselaer Polytechnic Institute, 110 8th St, Troy, New York, USA, 12180; John S. Erickson, erickj4@rpi.edu, Rensselaer Polytechnic Institute, 110 8th St, Troy, New York, USA, 12180; Oshani Seneviratne, senevo@rpi.edu, Rensselaer Polytechnic Institute, 110 8th St, Troy, New York, USA, 12180; Kristin P. Bennett, bennek@rpi.edu, Rensselaer Polytechnic Institute, 110 8th St, Troy, New York, USA, 12180.

Manuscript submitted to ACM

a decentralized way. Banking has been remade into **lending protocols**, currency exchanges have been recreated as **decentralized exchanges**, etc. Each DeFi protocol comes with its own set of novelties, attempting to innovate on traditional financial infrastructure in a unique way. As DeFi grows and matures, creating tools that help track and interpret changes in the DeFi ecosystem should prove interesting and useful, as the richness and volume of data publicly available in DeFi is unprecedented. We propose survival analysis tools which help address two big-picture questions about DeFi: *"How are users behaving within DeFi protocols?"* and *"How are user-level behaviors within DeFi protocols changing through time?"*

For this work, we demonstrate our approach on Aave, one of the largest lending protocols [4][1]. Our general survival analysis approach, however, can be readily generalized to the many existing and new DeFi protocols. Aave [4] has a market size of over $8.14 billion as of November 1, 2022, and is deployed on seven different networks including Ethereum, Polygon, and Avalanche, to name a few. Lending protocols serve a similar role in DeFi to banks in traditional finance; they allow users to deposit cryptocurrencies into a lending pool, granting them interest on their deposited assets. Users can redeem the deposits as desired. Users can also borrow funds from the lending pool so long as they provide sufficient collateral for the asset type and quantity they seek to borrow. Users can repay loans or keep them as long as desired as long as they maintain sufficient collateral. Should the value of their collateral drop too much or they allow too much interest to accrue on their loans, their account is at risk of being liquidated. Liquidations in Aave are one of the methods implemented by the developers to allow for the maintenance of the health of the lending pool. The protocol is configured to pay out a small bonus to any user (called a keeper) who is willing to repay another user's unhealthy loan balance in exchange for the collateral. When the keeper repays a liquidatee's loan, they receive the appropriate proportion of the liquidatee's collateral (their deposited assets), and also receive a small, percentage-based reward from Aave in order to incentivize liquidation.

Since these transactions are the building blocks of DeFi data streams, we applied a method of data-driven time-to-event analysis called "survival analysis" to the transactions streams. Survival analysis in the context of DeFi proves a versatile technique for derivation of myriad insights into macro-level user behaviors and how these behaviors have changed through time. We show here just a few of the many compelling results from applying survival analysis methods to Aave's transactions. This paper is organized as follows: in the methods section, we describe the source and structure of Aave data used, the survival analysis methods employed, and an overview of the application created in conjunction with this paper. In the results section, we demonstrate the use provided by survival analysis to answer various questions. We conclude with a discussion of the contributions of this work and promising directions for future work.

2 METHODS

2.1 Transaction-Level Data

The data used for this analysis is from The Graph[2]. Aave pushes its own data to The Graph, and each network Aave is deployed on has its own subgraph. These subgraphs are structured identically, at least with respect to the transaction-level data with which our analyses are concerned. We have collected all of the transaction data from nine subgraphs maintained by Aave, which include the data from the following Aave markets: Ethereum [6] (both Mainnet and Automated Market Maker versions), Polygon [11], Avalanche [22], Optimism [28], Harmony [25], Fantom [7], and Arbitrum [12]. There are currently two versions of Aave in deployment, referred to simply as V2 [5] and V3 [10] (V1 went by a different name and is, for all intents and purposes, no longer in use). The Ethereum markets only operate

[1] https://aave.com/
[2] https://thegraph.com/en/

on AaveV2, the Polygon and Avalanche markets have both AaveV2 and AaveV3 versions in deployment, and the rest are exclusive to AaveV3. We illustrate survival analyses of a single market and across different markets. Since, the various Aave markets come with the advantages and disadvantages of the networks (blockchains) on which they are deployed, examining differences in user behaviors across markets can naturally lead to interesting insights regarding how user behaviors change across markets. The data used here was collected from the time of AaveV2's deployment on November 30, 2020 through October 1, 2022.

The majority of the analyses presented in this paper focus solely on the Aave Ethereum market. This is for a few reasons: Aave is still fairly new, and the Ethereum market is the oldest of the deployments which means that its data spans the longest duration; the value locked on the Ethereum market is, as of November 1, 2022, approximately $6.16 billion, accounting for nearly 75% of the value locked across all of Aave's markets; and, the number of transactions from this market comprises only roughly 5% of the total transactions across all markets combined which makes computation of its associated survival data easier to test. For the data used in this paper, we include five transaction types from Aave: deposits, redeems, borrows, repays, and liquidations. There are some other transaction types that Aave collects, but they vary from market to market and the quantity of these transactions are so low that we do not feel they are important for the sake of this analysis; e.g., swaps (a user swapping from stable to variable borrow rates, or vice versa,) comprise just 3,382 transactions out of the 1,475,175 in Aave's Ethereum market, and is not tracked in other markets. A summary of the data used from the Ethereum market is given in table 1. Descriptions of each transaction type are provided in section 2.1.1 below, and table 2 shows a sample of the raw transaction data.

Note that Aave allows the usage of 92 different coins across all markets. For insightful analysis, we divide the coins into two types: stablecoins and non-stablecoins. A stablecoin is from a class of cryptocurrencies that attempts to offer price stability, typically in terms of USD. The other types of coins in the dataset are non-stablecoins. A breakdown of the coins available in Aave's Ethereum market and whether they are classified as stable or non-stable can be found in Appendix B. There are 11 different stablecoins in the Aave's Ethereum market, and 28 other coins which we call non-stable.

Table 1. Summary of transaction types from Aave's Ethereum Market collected from November 30, 2020 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	191,103	$350,886.90	$11,048.17
Deposit	363,282	$653,439.90	$11,100.46
Redeem	260,079	$871,974.2	$25,044.53
Repay	132,567	$481,908.4	$22,352.36

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	26,040	$39,026.36	$41,571.71

2.1.1 Transaction Type Descriptions.

The most fundamental transaction type in Aave is the "deposit". Like opening a savings account with a bank, deposited assets will slowly accrue interest, providing users some incentive to use the platform even passively. Naturally, users can withdraw their deposited assets from the protocol. The transaction type that effectively mimics withdrawing money from an account is called "redeem."

Table 2. Summary of transaction-level data from Aave's Ethereum Market

Date and Time	Type	User	Coin Symbol	Amount	Amount (in USD)	⋯
December 1, 2020 05:15:00	Deposit	<ID>	USDT	100.00	100.00	⋯
December 1, 2020 05:15:30	Borrow	<ID>	XSUSHI	15.52	100.00	⋯
⋮	⋮	⋮	⋮	⋮	⋮	⋱
September 30, 2022 23:50:00	Repay	<ID>	DAI	25,000.667	24,978.34	⋯
September 30, 2022 23:50:45	Redeem	<ID>	WETH	3.652	8,976.09	⋯

An important characteristic of an account in Aave is how much collateral a user has posted. When a user deposits a currency into Aave, in most cases they are able to select whether they want that currency to be used as collateral in their account (certain currencies are disallowed as collateral by Aave). Having one or more assets enabled as collateral in an account will allow a user to "borrow" assets from the Aave lending pool. The amount of cryptocurrency a user wants to borrow is capped by a percentage of the total value of the assets they have enabled as collateral in their account. These assets are all valued relative to their conversion rate with the Ethereum cryptocurrency. When a loan has been taken out through a borrow transaction, the user can "repay" the loan over time.

The most complex transaction type in Aave is the "liquidation." When someone borrows funds in Aave, they borrow them against the value of their collateral assets. As the loan accrues interest, and as the values of the collateral and borrowed currencies fluctuate with the market, a loan that was originally healthy can become unhealthy. If the health of a user's account (calculated using the relative value of the user's collateral assets and borrowed assets) deteriorates too much, the user's account can be liquidated. This means another user (called a keeper) is allowed to pay off a portion of the unhealthy loan and claim the appropriate portion of the user's collateral, and gain a small bonus from the Aave protocol for having done so. Liquidations are an important method for keeping the protocol healthy overall, since they provide an incentive for someone to repay a loan that was otherwise losing value for the lending pool.

2.2 Creating Survival Analysis Data

Survival analysis is a collection of statistical procedures for data analysis in which the variable of interest is the time from an index event until the outcome event [8]. Basically, survival analysis allows for time-to-event analysis. There are four primary choices to make when using survival analysis: an "index event," an "outcome event," a "covariate" of interest, and an "observation period." Index events trigger the beginning of a single observation. After some time has elapsed following an index event, an outcome event may occur. Outcome events trigger the end of an observation that began with an associated index event and the appropriate elapsed time is calculated. It is possible as well that the period of observation ends prior to the occurrence of an outcome event for a specific index event. In this case, the event is considered to be "right-censored," and a note of this censoring is made. The elapsed time is the time between the index event and the end of the observation period. The survival data will be segmented by each possible value of the chosen covariate, allowing us to see the effects the covariate has on the outcomes.

As an example, consider the question "How long do users take to repay stablecoins versus non-stablecoins after a borrow?" We can use survival analysis to answer this question by choosing borrow transactions as index events, repay transactions of the same coin (i.e., the principal reserve coin) as outcome events, the coin type (stable or nonstable) as the covariate, and an observation period of interest. For the sake of this example, let the observation period be the time from Aave's inception until the end of our data collection (Oct. 1, 2022). From these choices, the survival-style data we need to create will look like table 3.

Survival analysis with respect to covariates can allows us to address different questions. Covariates can be any factor associated with an index event. For instance, in the example in table 3, we added the covariates such as which currency was borrowed (i.e., coin), or whether the currency is a stablecoin (i.e., coin type.) New covariates can easily be defined and added as desired to achieve analysis goals.

Table 3. Sample survival-style data with **index events** being borrows and **outcome events** being repayments, with additional covariates.

Elapsed Time (days) from Index Event	Censored?	User	Coin	Coin Type	Aave Market	⋯
12.6	False	<ID>	WETH	Non-Stable	Ethereum	⋯
68.3	False	<ID>	USDC	Stable	PolygonV2	⋯
54.1	True	<ID>	DAI	Stable	Optimism	⋯
⋮	⋮	⋮	⋮	⋮	⋮	⋱
155.0	False	<ID>	AMPL	Non-Stable	AvalancheV2	⋯

2.2.1 *Computing Survival Data with Different Observation Periods.*

The choice of observation period allows for additional versatility in what questions we can answer with survival analysis, and brings with it slightly different steps for computing the survival data. For the work presented here, we use two different observation periods. The default observation period we use is **since inception**. For survival data computed since inception, we consider data from the start of Aave through the end of our collected data (October 1, 2022). This observation period only requires consideration of right-censored events, where the index event is observed but no outcome event is observed. In this case, we mark that an index event occurred and that a duration from the time of the index event until the end of the observation period has passed, but that the event is censored. Unless otherwise specified, this is the observation period used in our analysis.

In section 3.4, we change how we handle the observation period to **quarterly**. Instead of having just one observation period for the entire history of Aave, we break the data into quarterly observation periods (01-Jan through 31-Mar, 01-Apr through 30-Jun, 01-Jul through 30-Sep, and 01-Oct through 31-Dec). We only include quarters that we have complete data for, which includes 2021 Q1 through 2022 Q3. We treat each quarter as a separate observation period, which means that it is possible for an outcome event to occur with no associated index event having taken place in the observation period. This leads to "left-censored" events. If an outcome event occurs, say, 30 days after the start of the observation period and it was not preceded by an associated index event, we record in the survival data that an observation occurred with 30 days of elapsed time, and that the event was censored.

2.3 Survival Analysis Methods

We use Kaplan–Meier survival curves [13] to visualize our survival data. To quantify the characteristics of the survival data we use a couple of different metrics: the Cox coefficients [9] with corresponding p-values; and the restricted mean survival time (RMST) [21]. We briefly explain how to read and interpret these below.

To create the Kaplan–Meier curves, compute the RMST, and calculate the Cox coefficients, we employ the `survminer` package [14] and the `survival` package [27][26] in R [19]. For Kaplan–Meier curves we use the `ggsurvplot` function from survminer package. To compute the RMSTs we use the survival package's survmean function. For calculating the Cox coefficients (and p-values), we use the coxph function from the survminer package.

2.3.1 *Kaplan–Meier Survival Curves.*

Kaplan–Meier curves show the probability of the outcome event having *not yet occurred* after an increasing amount of time has elapsed. Interpreting a Kaplan–Meier curve requires the understanding of three main components: the index and outcome events used to create the underlying survival data, and the covariate selected as a strata. Numerous Kaplan–Meier plots can be created from the same survival data based on which covariate is selected for the strata, and for each unique value of the selected covariate, a separate Kaplan–Meier curve will be drawn on the same plot. To read the Kaplan–Meier plots, then, we will use figure 1(a) as an example.

In figure 1(a), the index events are "when a user borrows from Aave", the outcome events are "when the same user next makes a repayment of the same coin that was borrowed", and the covariate selected as the strata is the principal coin type. Since there are two choices for coin type (stable and non-stable), there are two Kaplan–Meier curves included in this plot. The individual curves show how likely it is for the outcome event to have not yet occurred after time passes. In this example, if we focus on the red curve (non-stable coin type), we see that 200 days after a borrow has been made, there is only a small likelihood of about 10% that the user has not made any repayments on that loan. Considering the curves for both stable and non-stable coins, we see that the borrows of non-stable coins have consistently lower probabilities of not having been repaid than the borrows of stablecoins. This means that users typically begin repaying non-stablecoin borrows faster than stablecoin borrows.

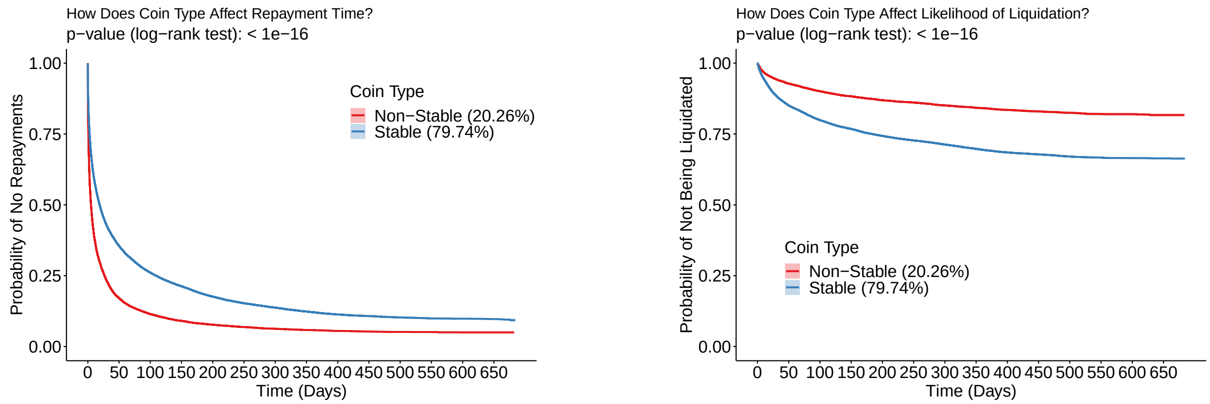
The statistical significance of the difference between any Kaplan–Meier curves in a plot are calculated and given as a p-value using the log-rank test. The p-values have been included atop the figures, and due to the magnitude of data present in the creation of these plots, these p-values tend to be very small. If the p-value is less than $1e - 16$, then we round it to 0. Lastly, each Kaplan–Meier curve is plotted with a 95% confidence interval included. Similar to the p-values rounding to 0, the confidence intervals tend to be so closely fit to the curves themselves that they often are not visible, but nonetheless they are included in each Kaplan–Meier curve drawn.

2.3.2 *Cox Proportional Hazards Regression.*

The Cox proportional-hazards model [9] is a model for quantifying the effects of covariates on the survival time of events. For a categorical covariate, Cox regression will use one of the values of that covariate as a reference point against which to compare the likelihood of survival through time. We compute and give the Cox coefficients for the non-reference values of the covariates. Cox coefficients that are positive indicate that the value of the covariate "reduces survival likelihood" (the outcome event is more likely to occur) relative to the reference. Negative Cox coefficients indicate the value of the covariate "increases survival time" (the outcome is less likely to occur). The larger in magnitude the coefficient is, the more significant the effect of that value of the covariate. We also give the p-value provided for each Cox coefficient to indicate the statistical significance of these results. The p-values are computed by a Wald test [29].

2.3.3 *Restricted Mean Survival Times.*

We use the restricted mean survival time to calculate the mean survival time for types of events stratified by covariates across the entire observation period. For instance, in table 5 we calculate the RMST of borrow-to-repay survival data



(a) Kaplan–Meier plot using borrows as index events, repays as outcome events, and principal coin type as covariates

(b) Kaplan–Meier plot using borrows as index events, liquidations as outcome events, and principal coin type as covariates

Fig. 1. Figure 1(a) shows that users make their first repayment for non-stablecoin loans significantly sooner than for stablecoin loans. Figure 1(b) shows that user accounts get liquidated more quickly following the borrow of a stablecoin than of a non-stablecoin. Both cases were statistically significant, with the log-rank test yielding p-values of 0.

stratified by coin type. For non-stablecoin loans the RMST is 58.73 days, meaning that over the course of the entire observation period the average time it took for users to repay loans of non-stable coins was 58.73 days. Likewise, for stablecoin loans the RMST is significantly higher at 117.97 days.

3 RESULTS

3.1 Insights Into User Behavior After Depositing Currency

We show that survival analysis can provide a useful picture of how users behave after depositing money into their accounts. Deposits are the natural first transaction for a user to make in the Aave lending protocol, since before borrowing they must deposit collateral. Thus, looking at how users behave after making deposits seems a natural place to begin our analysis. Using deposits as index events, the first subsequent transaction of any type as outcome events, and the outcome event type as the covariate, we get figure 2. Note that by using the outcome type as the covariate, we are limiting our observations to those that did witness an outcome event. Of the 363,282 deposits, there were 28,519 that were not followed by another transaction and thus are not represented in this analysis.

In figure 2, provides us insights into what users are doing with their deposits. Users frequently make deposits and then immediately use them as collateral in borrows (32.65% of next transactions after a deposit are borrows). They tend to do so more quickly than any other transaction. However, based on the percentages of each outcome event, users are more likely to make an additional deposit (36.13%) as their next transaction following a deposit fairly rapidly. User depositing and then using the funds to repay loans slightly less rapidly, but are less common (8.19%), deposits followed by redeems occurs 22.27% of the time, but it tends to take much longer than the time to all other outcome types except liquidations. Not surprisingly since user have to borrow before making a liquidation, deposit to liquidations occur less frequently and are tend to take the longest.

We can use table 4 to better understand this analysis. From the RMSTs, we see that when a user borrows after a deposit, the mean time to do so is only 2.18 days. This is by far the action that users are the fastest to make. The next fastest action is making another deposit, which the mean time to make is 6.89 days. Users are also more quick to make repayments after deposits than they are to redeem their deposits. This seems intuitive enough, since a user will need to make a deposit of whatever currency they had borrowed in the past in order to repay it. The Cox coefficients and significant p-values further support these conclusions.

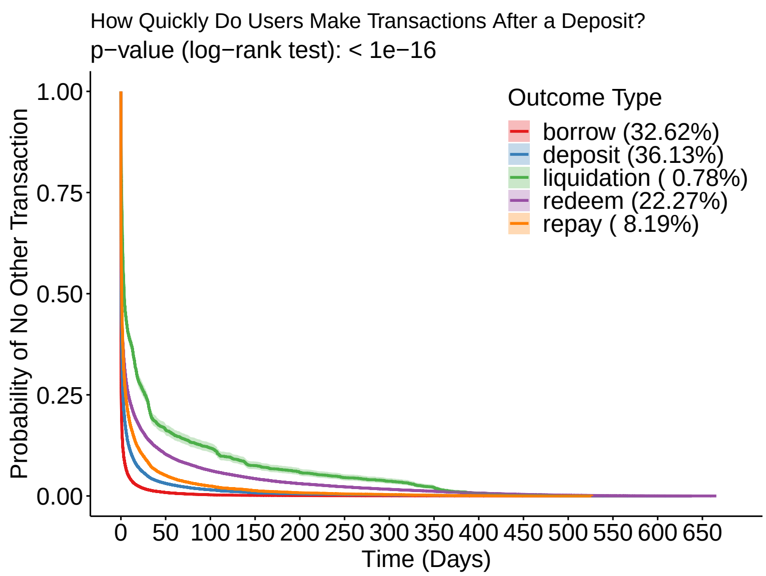


Fig. 2. Kaplan–Meier plot using deposits as index events, next transaction as outcome events, and type of outcome transaction as covariates. Legend include percentage of each transaction type after borrow.

Table 4. Deposit-to-Next-Transaction: Quantifiers of the survival data corresponding to figure 2, using deposits as index events, next transaction as outcome events, and type of outcome transaction as covariates.

Outcome Type	Percentage	RMST (days)	Cox Coefficient	P-value
borrow	32.62	2.18	Reference	Reference
deposit	36.13	6.89	-0.50217	<1e-16
liquidation	0.78	36.62	-1.25017	<1e-16
redeem	22.27	21.66	-0.87077	<1e-16
repay	8.19	10.64	-0.77642	<1e-16

3.2 Insights Into Loans and Their Outcomes

Next we consider user behavior with respect to loans. Enabling users to take out loans without the need for explicit approval by another party is one of the more unique features of DeFi lending protocols. A user is able to borrow currency based on the amount of other currencies they have deposited into the lending pool and are willing to post as collateral for their loan. This smart-contract enabled lending is interesting, because it has no direct analog in traditional banking and also because the reasons for which someone would take out a loan in this setting are not immediately obvious. The requisite over-collateralization for loans makes motivation for loans less obvious, since to be able to take out a loan requires the prior ownership of a relatively greater amount of crypto-assets in the first place. Since the allowance of lending is the primary reason for lending protocols to exist, we feel that emphasizing behavior related to loans and their outcomes is the most natural and interesting place for survival analysis.

In Aave, loans are indefinite until some combination of two potential outcomes occur: repayment of the loan and liquidation of the account. Loans can be repaid in as many or as few repay transactions as the debtor desires, and there is no required timeline to pay back loans; the loan can last for as long as the user's account remains healthy. If the user's account becomes unhealthy, they can be liquidated by another user. When a user is getting liquidated, the keeper can choose one of the unhealthy user's collateral coins and one of their principal reserve coins and pay off up to 50% of that principal coin the user has borrowed in order to claim an equal proportion of the collateral coin. The liquidator is paid a small bonus by Aave to incentivize this action. This is the other way that a loan can be partially repaid.

In either case, we will be using borrow transactions as index events and either repay transactions or liquidation transactions as the outcome events. Note that these outcome events only mark the first repayment or the first liquidation following the borrow, not the time for the loan to be fully repaid or fully liquidated.

3.2.1 *How does coin type affect the outcomes of loans?*

One covariate we hypothesize to have a significant effect on how users behave with their loans is whether the loan is of a stable or non-stable coin. We compute survival data using borrows as index events and either repays (figure 1a) or liquidations (figure 1(b)) as outcome events, using coin type as covariates. For repayments, we see users repaying non-stable coins significantly quicker than stable coins. From table 5 we get that the RMST for repaying non-stable coins is 58.73 days, which is just about half as long as the RMST for repaying stable coins (117.97 days). In contrast, non-stable coins are liquidated much less frequently than stable coins. The mean time it takes for stablecoin loans to be liquidated is 496.17 days, which is almost 100 days sooner than the mean time of 585.79 days for non-stablecoin loans.

Table 5. Borrow-to-Repay: Quantifiers of the survival data corresponding to figure 1(a), using borrows as index events, repays as outcome events, and coin type as covariates.

Coin Type	Percentage	RMST (days)	Cox Coefficient	P-value
Non-Stable	20.26	58.73	Reference	Reference
Stable	79.74	117.97	-0.49970	<1e-16

Table 6. Borrow-to-Liquidation: Quantifiers of the survival data corresponding to figure 1(b), using borrows as index events, liquidations as outcome events, and coin type as covariates.

Coin Type	Percentage	RMST (days)	Cox Coefficient	P-value
Non-Stable	20.26	585.79	Reference	Reference
Stable	79.74	496.17	0.73398	<1e-16

3.2.2 *How does loan size affect the length of loans?*

Next we examine the effect of loan size on the time it takes for users to pay off the loan. In traditional finance, the length of loan is specified in a contract and we tend to expect loans to take longer to pay off the larger they are. For example large home mortgages are longer than smaller car loans. In Aave, loans do not have defined lengths, typically the collateral and loan principal are in different coins, and the interest rates can vary drastically throughout the course of a loan. Despite these differences, we may still expect larger loans to be repaid over a longer period of time in Aave.

Surprisingly, we see the exact opposite behavior in DeFi. In Aave's history, the larger a loan is, then the faster the loan is likely to be fully repaid. We see this result in figure 3, and the fact that the differences are significant is clearly shown by the large separation between the curves. The p-value a for the log-rank test confirms this significance. In table 7, we give the RMST and Cox coefficients for the different quartiles of loan amounts. The smallest quarter of loans have a mean survival time of 493.54 days, meaning that users are averaging close to 500 days to fully repay the smallest loans in Aave. In contrast, the largest loans are being repaid in just 166.26 days on average. The Cox coefficients confirm that the larger a loan is, the higher the risk is for the loan to be fully repaid through time.

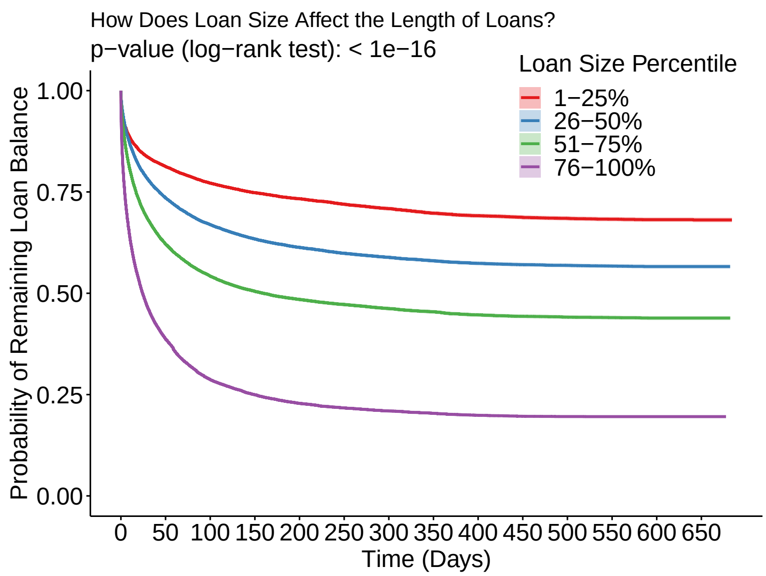


Fig. 3. Kaplan–Meier plot using borrows as index events, the full repayment of the borrow as outcome events, and loan size quartiles as covariates.

Table 7. Borrow-to-Full-Repayment: Quantifiers of the survival data corresponding to figure 3a, using borrows as index events, full loan repayments as outcomes, and loan size as covariates.

USD Amount Quartile	Percentage	RMST (days)	Cox Coefficient	P-value
1st Quartile (1-25%)	25.13	493.54	Reference	Reference
2nd Quartile (26-50%)	25.03	418.66	0.40379	<1e-16
3rd Quartile (51-75%)	25.00	333.94	0.78995	<1e-16
4th Quartile (75-100%)	24.84	166.26	1.49488	<1e-16

3.2.3 How do coin types influence risk of liquidation?

We hypothesize that the combination of stable and non-stable coins of the principal reserve and the collateral may lead to further insight into the risk of borrows. As shown in 4, we stratify the borrow-to-liquidation data by factoring in what collateral was purchased and what principal types were specifically paid off by a liquidator. Since we are splitting the curves by what principal and collateral were paid off and purchased at the time of the liquidation, all the curves do end up with a 0% probability of survival, similar to the curves in figure 2. Again though, we can still use the curves to gain insight into the relative riskiness of the principal:collateral combinations that people can have in their accounts. According to the log-rank test, the differences in the curves are statistically significant.

The definition of the outcome event in this analysis is quite different. We aggregated user's liquidation events to gain more information as to which coins the users have used as collateral in their account. Even though each liquidation transaction only records one principal type and one collateral type, sometimes a user will be the subject of multiple liquidations in quick succession. It would be inaccurate to consider these liquidations as separate events; they really are all part of one bigger liquidation event. Thus, in our transaction data, if a user is liquidated multiple times in quick succession with no intermittent non-liquidation transactions, we aggregate them into one bigger liquidation transaction. The outcome event is the combined liquidation transaction, with the time being the first liquidation transaction. This lets us see whether there were multiple types of collateral and principal coins involved in the event. Thus, if a user has both stablecoins and non-stablecoins in their account as collateral, or if they've taken out loans of both stablecoins and non-stablecoins, we mark the collateral or principal, respectively, as "Mixed."

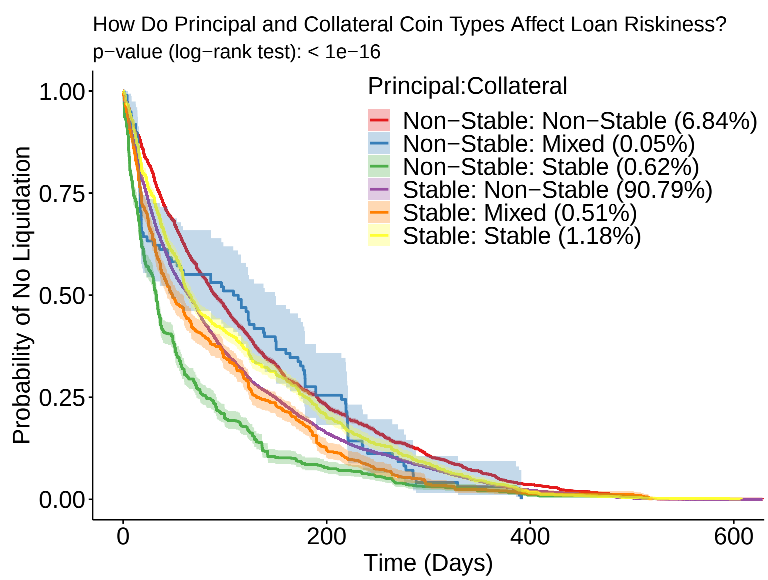


Fig. 4. Kaplan–Meier survival curves using borrows as index events and liquidations as outcomes using different combinations of principal and collateral coin types as covariates.

Using the stable:stable combination of principal and collateral coins as the benchmark, we get the quantification of risk via seen in table 8.

The results show that the combination of non-stablecoin principal with non-stablecoin collateral tends to be safer than that of stablecoin principal and collateral. This is a little surprising, since the relative value of an account's principal and collateral are what causes its ability to be liquidated. When functioning properly, stablecoins should always have the same relative value and the only reason an account should liquidate with stablecoins as both the principal and the

Table 8. Borrow-to-Liquidation: Quantifiers of the survival data corresponding to figure 4, using borrows as index events, account liquidations as outcomes, and the combination of principal and collateral coin types as covariates.

Principal:Collateral Combination	Percentage	RMST (days)	Cox Coefficient	P-Value
Stable:Stable	1.18	113.63	Reference	Reference
Non-Stable:Non-Stable	6.84	128.28	-0.11877	1.22094e-07
Non-Stable:Mixed	0.05	117.73	-0.00409	0.96836
Non-Stable:Stable	0.62	64.56	0.51160	< 1e-16
Stable:Non-Stable	90.79	102.61	0.09348	7.47579e-06
Stable:Mixed	0.51	90.54	0.20414	5.75758e-08

collateral would be when enough interest has accrued and gone unpaid on the account. Non-stablecoins' values are much more volatile, and if the relative value of the principal coin is not in sync with the value of the collateral coin, the account could be liquidated much more quickly. However, this is not what we see. This likely indicates that users who borrow non-stablecoins using non-stablecoins as collateral are using collateral assets that are less likely to drop in value than their principal asset(s).

We also see that loans are liquidated more quickly when they have non-stablecoin principal and stable collateral. This behavior makes sense logically during a bull market when many coins will be gaining value relative to the USD which most stablecoins are pegged to. If the value of the principal assets rise relative to the value of the collateral, the loan can become unhealthy very quickly. If an account consists of stablecoin principal and non-stablecoin collateral, we see a slight increase in risk relative to stable:stable accounts. This combination also accounts for the vast majority (90.79%) of liquidated accounts.

3.3 Insights Into Differences Between Markets

The analysis so far has focused on Aave's Ethereum market. We now examine how user behavior compares in Aave's other markets of AaveV3. The transaction volumes for the other markets are given in Appendix A. The six V3 markets were launched on March 12, 2022 and the data used here runs through October 01, 2022. Because we only have 203 days of data for the V3 markets, we restrict the time window when computing the RMSTs to the first 203 days of observation.

In table 9, we compute survival data using borrows as index events, liquidations as outcome events, and the Aave market as covariates. The RMSTs and Cox coefficients paint an interesting picture here. Using the Ethereum market as a benchmark, we see all but one of the V3 markets showing less risk of liquidation following a borrow. The Fantom market, which accounts for just 1.87% of the borrow events in these transactions, has a RMST of 145.72 days. This indicates that in the first 203 days of these markets, the mean time it takes for a loan to be liquidated following a borrow is 23.02 days less than the mean time for borrows to be liquidated in the Ethereum market. In contrast, the other V3 markets show significantly longer times to liquidation. For instance, the Arbitrum market shows a mean survival time of 193.63 days, which is almost a 25 day increase over the Ethereum market's survival time.

In table 10, we see a different side of the markets. Using deposits as index events, borrows as outcome events, and again using the Aave market as covariates, we can see how quickly users will borrow funds after making a deposit in Aave. In the Ethereum market the mean time for users to borrow after a deposit is 85.76 days, and again we see the biggest contrast in the Fantom market, where the mean time is only 44.55 days. The Polygon market, which is the market with the second-biggest share of deposits in this data at 23.76% of the deposits, also shows users being much more likely to borrow funds following a deposit. The mean survival time for the Polygon market is just 58.63 days.

One factor that likely causes differences in user behavior across markets is the size of transaction fees in the market. The transaction fees are just flat fees, not scaled by transaction size, so in markets with higher transaction fees, making smaller transactions is more heavily penalized. Additionally, making lots of transactions is more heavily penalized. While our data source does not tell us how much users paid in transaction fees, we do know that the Ethereum market has non-negligible fees, usually at least $10. The other most popular markets of Polygon and Optimism have very low transaction fees, usually costing less than $0.01. As a result, we suspect that these markets attract retail users who are looking to experiment more with the protocol and engage in behaviors such as yield farming to make shorter term profits, whereas the Ethereum market likely attracts much larger, institutional users.

Table 9. Borrow-to-Liquidation: Quantifiers of survival data using borrows as index events, liquidations as outcome events, and the Aave market where the transactions took place as covariates. The RMSTs were computed with a cutoff of 203 days to account for the shorter span of data from the non-Ethereum markets.

Market	Percentage	RMST (days)	Cox Coefficient	P-Value
Ethereum	30.86	168.74	Reference	Reference
Arbitrum	7.42	193.63	-1.37313	< 1e-16
Avalanche	11.99	185.93	-0.76679	< 1e-16
Fantom	1.87	145.72	0.55972	< 1e-16
Harmony	0.66	179.06	-0.468234	< 1e-16
Optimism	14.33	191.90	-1.24776	< 1e-16
Polygon	32.86	184.03	-0.63496	< 1e-16

Table 10. Deposit-to-Borrow: Quantifiers of survival data using deposits as index events, borrows as outcome events, and the Aave market where the transactions took place as covariates. The RMSTs were computed with a cutoff of 203 days to account for the shorter span of data from the non-Ethereum markets.

Market	Percentage	RMST (days)	Cox Coefficient	P-Value
Ethereum	29.97	85.76	Reference	Reference
Arbitrum	9.86	89.05	-0.01895	1.66912e-05
Avalanche	14.52	52.72	0.43877	< 1e-16
Fantom	2.12	44.55	0.57727	< 1e-16
Harmony	1.52	94.88	-0.13866	< 1e-16
Optimism	18.26	75.31	0.23583	< 1e-16
Polygon	23.76	58.63	0.39877	< 1e-16

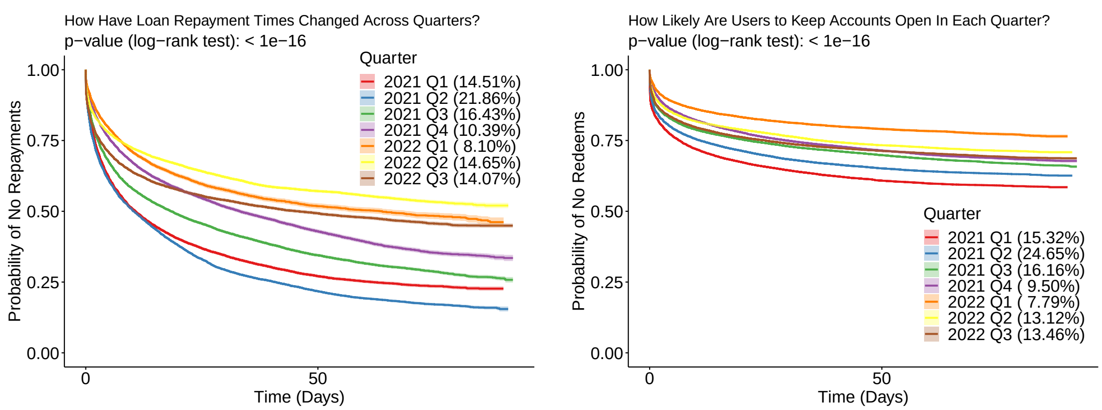
3.4 Insights Into Evolving Behavior

Since DeFi is still in its infancy, we would expect to see changes in macro-level user behaviors as people get excited about the new technology and start using it, but perhaps do not have a well-established understanding for how the technology should be used. In this section we look at a couple of results of how user behaviors have changed from quarter to quarter. We again compute how long it takes for users to fully repay loans after borrowing money (figure 5(a)), and we also look at behavior relating to how long users keep funds in their accounts (figure 5(b)).

In figure 5(a), we look at Kaplan–Meier curves for borrow-to-repay events with quarters as covariates. Each distinct curve represents a different quarter of transactions. We note that, while the shapes of the curves are consistent from quarter to quarter, there are drastic differences between how quickly users are repaying loans depending on the quarter.

For instance, we can see that in 2022 Q2, in the entire quarter only about 50% of the loans were paid back. This is in contrast to 2021 Q2 where over 75% of the loans were repaid by the end of the quarter. The Cox coefficients for these curves are in table 11. What we can observe from these coefficients is that the tendency through time has been for users to take longer to repay their loans. 2021 Q2 showed a slightly increased risk of loans being repaid during the quarter, but the subsequent four quarters each show reduced likelihood of loans being repaid. This trend was broken in 2022 Q3, but the fact remains that users have been tending to slow down their repayment schedules as Aave has matured.

In figure 5(b), we change our focus to how long users are leaving money in their Aave accounts. We hypothesize that users who leave money in their accounts longer are more invested in both the protocol of Aave and in the emerging DeFi ecosystem, and thus it is interesting to see how the duration that users keep money in their accounts has changed through time. Similar to the behavioral changes seen in figure 5(a), the time it takes for users to redeem funds after depositing them changes drastically from quarter to quarter. Generally speaking, users are leaving funds in their accounts much longer than the terms of loans, but the relative difference from quarter to quarter for how quickly users are withdrawing funds is significant. We see a similar trend in the deposit-to-redeem analysis as in the borrow-to-repay analysis, where the first quarter in the data shows users being the most likely to redeem funds from their accounts, and the general trend has been that as time has passed, the risk of users withdrawing funds has decreased. Again, these observations are confirmed by the Cox coefficients in table 12.



(a) Kaplan–Meier plot for borrow-to-repayment

(b) Kaplan–Meier plot for deposit-to-redeem

Fig. 5. Kaplan–Meier Survival Curves showing the differences in repayment schedules and withdrawal schedules throughout individual quarters since Aave V2 launched. 2020 Q4 has been removed due to the small portion of the quarter for which Aave existed (November 30-December 31, 2020).

3.5 Survival Analysis App

Since there are so many choices for index events, outcome events, and covariates, we thought it appropriate to help create these analyses and allow other people to explore our survival data. The app can be accessed at this link: https://inciteprojects.idea.rpi.edu/defitoolkit/app/defitoolkit/. This app allows the user to create Kaplan–Meier survival curves and tables of related data similar to what has been presented in this paper. A sidebar on the left of app allows the user to select the parameters for the analysis, including which DeFi protocol to pull data from, which version of the protocol to use, which market to look at, whether the analysis should be computed in totality or quarterly as described

Table 11. Borrow-to-Repayment: Quantifiers of the survival data corresponding to figure 5(a), using borrows as index events, first repayments as outcome events, and the quarter in which the events took place as covariates. This data was computed using each quarter as a separate observation period.

Quarter	Percentage	RMST (days)	Cox Coefficient	P-Value
2021 Q1	14.51	30.91	Reference	Reference
2021 Q2	21.86	26.64	0.10415	< 1e-16
2021 Q3	16.43	36.93	-0.19548	< 1e-16
2021 Q4	10.39	44.36	-0.44371	< 1e-16
2022 Q1	8.10	52.06	-0.67579	< 1e-16
2022 Q2	14.65	55.77	-0.79187	< 1e-16
2022 Q3	14.07	48.61	-0.51226	< 1e-16

Table 12. Borrow-to-Redeem: Quantifiers of the survival data corresponding to figure 5(b), using deposits as index events, first redeems as outcome events, and the quarter in which the events took place as covariates. This data was computed using each quarter as a separate observation period.

Quarter	Percentage	RMST (days)	Cox Coefficient	P-value
2021 Q1	15.32	58.70	Reference	Reference
2021 Q2	24.65	62.36	-0.15160	< 1e-16
2021 Q3	16.16	66.07	-0.31816	< 1e-16
2021 Q4	9.50	67.94	-0.42282	< 1e-16
2022 Q1	7.79	74.23	-0.76368	< 1e-16
2022 Q2	13.12	69.43	-0.48187	< 1e-16
2022 Q3	13.46	67.43	-0.37813	< 1e-16

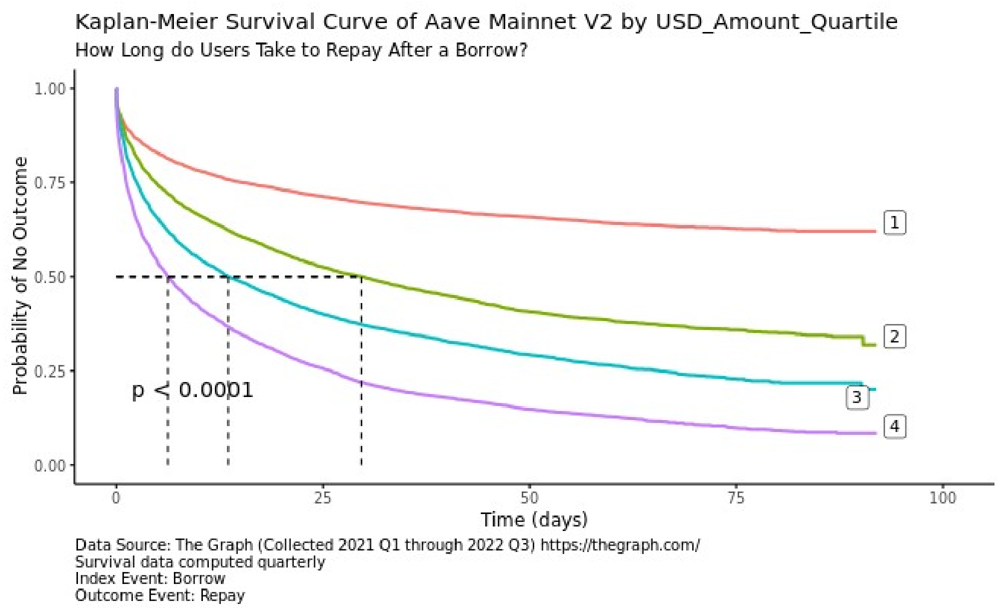
in section 2.2.1, the index and outcome events of interest, and a covariate of interest. As we continue to expand the scope of our work, more protocols and markets will be added to the app. At the time of writing this paper, we only have Aave data incorporated in the app.

With all the parameters set, the app generates the Kaplan–Meier curve and data table on the right. For example, the screenshot shown in figure 6 shows the analysis of Aave V2 Ethereum's market computing the survival data quarterly with borrows as index events, repays as outcome events, and the loan size quartile as the covariate.

4 RELATED WORK

With an over-collateralized loan, a borrower must post collateral which exceeds the value of the debt. This way, collateralization simultaneously ensures that the lender (likely a smart contract) can recover their loaned value. The "health factor" (HF) is a custom threshold in DeFi lending systems. If the debt collateral falls below the HF (typically below 1), the debt position may be opened for liquidation. Then the liquidators can purchase the locked collateral at a discount and close the borrower's debt position. Thus, leveraged positions are subject to liquidation when the debt becomes unhealthy, and a liquidator can repay the debt and benefit from a liquidation spread.

Given this novel form of automatic lending, a growing body of literature has studied liquidations on borrowing and lending platforms in DeFi. Qin et al. [18] have analyzed risk management provided by liquidators, acting on the protocol's user accounts. They have measured various risks that liquidation participants are exposed to on four major Ethereum lending pools (i.e., MakerDAO [?], Aave, Compound [20], and dYdX [1]) and quantified the instabilities of existing lending protocols. They have illustrated that the commonly used incentive mechanisms tend to favor liquidators



Descriptive Statistics of Aave Mainnet V2

USD_Amount_Quartile	Percentage	RMST	Hazard_Ratio	Cox_Coefficient	P_Value
1	26.3861	63.1151	Reference	Reference	Reference
2	24.6733	43.3401	1.9716	0.6789	0
3	24.0891	32.4622	2.7777	1.0216	0
4	24.8515	19.9717	4.1554	1.4244	0

Fig. 6. A screenshot of the Survival Analysis App showing a Kaplan–Meier curve and table of associated quantifiers. The analysis in this screenshot was created using data from the Aave V2 Ethereum market. The survival data was computed quarterly. The index and outcome events are borrows and repays, respectively. The covariate chosen was USD_Amount_Quartile, which is the same loan amount quartile used in section 3.2.2.

over borrowers, causing the problem of so-called over-liquidation, leading to unnecessary high losses for borrowers. The only recourse the borrowers have to avoid such liquidations is to monitor their loan-to-value ratio when the market changes quickly because even a random drop in market prices can result in a cascade of liquidations. If there are any drops in the market, it can lead to self-accelerating pressure to sell, which further causes more problems for a blockchain-based DeFi, such as network congestion that leads to steep gas costs. We witnessed such an event in the Ethereum market collapse of March 13, 2020[3] that left some borrowers unable to react, despite imminent liquidations.

[3]https://coinmarketcap.com/historical/20200313

It can be particularly bad for borrowers who get liquidated if market prices recover after a dip again, leaving them deprived of subsequent upward price participation. In general, regardless of market conditions, liquidations in DeFi are widely practiced, and related works such as Qin et al. [18] have quantified that over the years 2020 and 2021, liquidators realized a financial profit of over 800M USD while performing liquidations.

Stablecoins play a significant role in liquidations, as they have several characteristics that are directly tied to liquidation mechanics. For example, a user may take a loan with a stablecoin as collateral with the intent of holding the loan indefinitely. If the stablecoin collateral is accruing higher interest than the borrowed principal coin, this can lead to a form of passive income. Early empirical evidence on the stability of crypto-backed loans with stablecoins has been studied by Kozhan and Viswanath-Natraj [15]. They specifically focused on the price volatility in the MakerDao protocol, which introduced the world's first decentralized stablecoin called Dai that is soft-pegged to the US Dollar, i.e., it uses a collateralized debt position mechanism to keep the price stable with respect to the US Dollar. They have analyzed how collateral stability increases peg stability and found a positive relationship between collateral risk and the price volatility of the stablecoin Dai.

The efficiency of lending pool liquidations has been studied by Perez et al. [17], in which they introduced a lending pool state model that is instantiated with historical user transactions observable in the Compound[4] implementation deployed on Ethereum. Their model abstraction facilitates the observation of state effects of each interaction and investigates the latency of user liquidations following the under-collateralization of borrowing accounts. Similarly, Bartoletti et al. [2] provide an abstract formal state transition model of lending pools and prove fundamental behavioral properties, which had previously only been presented informally in the literature. Additionally, the authors examine attack vectors and risks, such as utilization attacks and interest-bearing derivative token risk.

As the demand for loans in crypto-assets grows, the borrowing interest rate goes up. In a bullish crypto market, speculators may be keen to borrow funds even if there is a high interest rate, in expectation of an appreciation in the assets of their leveraged long position as demonstrated by Xu et al. [30]. Such an environment is advantageous for lenders, resulting in higher yields to them. Compound and Aave, two major DeFi lending protocols, have witnessed the borrow APY of the stablecoin USDC increasing from a low of 0.84% in September 2022 to as high as 10% in April 2021 (as of this paper writing in November 2022, the APY is hovering at 1.1% in Compound and Aave, but other protocols such as Nexo offer upwards of 10% APY, and now-defunct protocols like Celcius and BlockFi used to offer similarly high APYs) [5]. In a bullish market, the yield generated is incorporated in interest-bearing tokens, such as `aTokens` from Aave analyzed in this paper. However, as was already noted, the wild fluctuations in the market result in unexpected liquidation events, as evidenced from this paper's results.

Most of the related works approach the issue of liquidation at a conceptual level or rely on aggregate flow data. In contrast, our paper uses transaction-level blockchain data to provide a more "microscopic" view of liquidations, combined with survival analysis techniques.

5 DISCUSSION AND FUTURE WORK

This work defines a pipeline for survival analysis of DeFi lending protocols which includes data aggregation, cleaning, converting to a data abstraction model, and performing powerful survival statistical analyses and visualizations to gain insights. Using Aave transaction data in the scenarios above, we have shown these survival analysis methods to be versatile tools for answering all kinds of questions in the DeFi sphere.

[4] https://compound.finance
[5] https://defirate.com/usdc/?amount=100&symbol=USDC&term=365&rate_type=lend

The possibilities for what questions to ask and answer with survival analysis are myriad. What we have presented in the results here were limited to just a handful of selections for index and outcome events, and using a few different variables to stratify the results. We showed that, counter to basic intuition, DeFi loans tend to be repaid more quickly when the loan is larger. We have shown how the combinations of stable and non-stable coins as both principal and collateral assets affect the riskiness of an account ending up being liquidated. Expanding the scope of the data briefly, we showed that depending on the Aave market, behaviors can change drastically. We hypothesize that these behavioral changes have mostly to do with the magnitude of the transaction fees in the market, and show that there are large differences in risk of liquidation over time between markets with high transaction fees and markets with low transaction fees. Lastly, we show that the data can be segmented into separate observation periods for each quarter, and that we can use these quarterly observation periods to find differences in user behavior through time.

This work represents just one step in the usage of survival analysis techniques to help form a complete picture of behavioral trends in the DeFi ecosystem. We note that these DeFi survival analysis techniques could be generalized to other DeFi lending protocols. DeFi survival analysis can be applied to any cryptocurrency protocol with transactions. Hazard analysis and other types of survival analysis and visualization methods could be used as well. In addition to this paper, we have also been preparing a toolkit for analysis of DeFi protocols through many different lenses. Survival analysis techniques are just one part of the toolkit, and eventually the results of these survival analysis applications will be integrated into other results to help form more robust analyses of DeFi protocols.

This research is that it only partially addresses the rich DeFi ecosystem, which has many interacting protocols and coin prices. We are already exploring the use of more advanced Artificial Intelligence (AI) methods for the analysis of transaction data developed for commerce and health [3, 24] that could incorporate more aspects of the DeFi ecosystem. These could be used for segmenting users and predicting behaviors and prices. Early results analyzing Aave transactions using Neural Temporal Point Processes are very promising [23]. DeFi represents an exciting new domain for AI research in transaction modeling since DeFi is a compelling use case, and all the datasets are public by definition.

Moving forward, we plan to continue using the survival analysis methods to help with ongoing work identifying and quantifying characteristics of the emerging DeFi ecosystem. We are working towards solidifying some methods for clustering the users of DeFi protocols and explaining what kinds of users each cluster represents. When the clusters are completed we plan to add user cluster as an additional covariate in our survival analysis, and we hope to see big differences in behavior across clusters. To improve the results relating to differences in user behavior over time, we plan to quantify the "bullishness" of the market through time and try to identify different behavioral trends based on how bullish the market is at the time. We also hope to incorporate flash loan data into our analyses as soon as we can correlate users' transactions between protocols, in order to get the true picture of how flash loans are being used.

All of this analysis will eventually be incorporated into our open-source DeFi Toolkit[6]. Currently the toolkit allows for the dynamic creation of survival analysis plots like what have been presented here. As we continue to refine certain techniques, they will be added to the toolkit to make for a more robust application.

ACKNOWLEDGMENTS

The authors acknowledge the support from NSF IUCRC CRAFT center research grant (CRAFT Grant #22003) for this research. The opinions expressed in this publication do not necessarily represent the views of NSF IUCRC CRAFT. This

[6]https://inciteprojects.idea.rpi.edu/defitoolkit/app/defitoolkit

work was supported by the Rensselaer Institute for Data Exploration and Applications (IDEA)[7]. We would like to thank Leyi Lin and Shuyu Meng for helping create and interpret results related to this work.

REFERENCES

[1] Antonio Juliano. 2017. *dYdX: A Standard for Decentralized Margin Trading and Derivatives.* Technical Report. https://whitepaper.dydx.exchange/

[2] Massimo Bartoletti, James Hsin-yu Chiang, and Alberto Lluch Lafuente. 2021. SoK: lending pools in decentralized finance. In *International Conference on Financial Cryptography and Data Security.* Springer, 553–578.

[3] Kristin P. Bennett. 2019. Artificial Intelligence for Public Health. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* 1–2. https://doi.org/10.1109/BIBM47256.2019.8983112

[4] Boado, Ernesto. 2020. *AAVE Protocol Whitepaper.* Technical Report. https://www.cryptocompare.com/media/38553941/aave_protocol_whitepaper_v1_0.pdf

[5] Boado, Ernesto. 2020. *AAVE Protocol Whitepaper V2.0.* Technical Report. https://cryptorating.eu/whitepapers/Aave/aave-v2-whitepaper.pdf

[6] Vitalik Buterin. 2014. *Ethereum: A next-generation smart contract and decentralized application platform.* https://github.com/ethereum/wiki/wiki/White-Paper

[7] Sang-Min Choi, Jiho Park, Quan Nguyen, and Andre Cronje. 2018. Fantom: A scalable framework for asynchronous distributed systems. https://doi.org/10.48550/ARXIV.1810.10360

[8] Taane G Clark, Michael J Bradburn, Sharon B Love, and Douglas G Altman. 2003. Survival analysis part I: basic concepts and first analyses. *British journal of cancer* 89, 2 (2003), 232–238.

[9] D. R. Cox. 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 34, 2 (1972), 187–220. http://www.jstor.org/stable/2985181

[10] Emilio Frangella, Lasse Herskind. 2022. *AAVE V3 Technical Paper.* Technical Report. https://github.com/aave/aave-v3-core/blob/master/techpaper/Aave_V3_Technical_Paper.pdf

[11] Anurag Arjun Jaynti Kanani, Sandeep Nailwal. 2020. Matic Network Whitepaper. https://github.com/maticnetwork/whitepaper.

[12] Harry A. Kalodner, Steven Goldfeder, Xiaoqi Chen, S. Matthew Weinberg, and Edward W. Felten. 2018. Arbitrum: Scalable, private smart contracts. In *USENIX Security Symposium.*

[13] E. L. Kaplan and Paul Meier. 1958. Nonparametric Estimation from Incomplete Observations. *J. Amer. Statist. Assoc.* 53, 282 (1958), 457–481. http://www.jstor.org/stable/2281868

[14] Alboukadel Kassambara, Marcin Kosinski, and Przemyslaw Biecek. 2021. *survminer: Drawing Survival Curves using 'ggplot2'.* https://CRAN.R-project.org/package=survminer R package version 0.4.9.

[15] Roman Kozhan and Ganesh Viswanath-Natraj. 2021. Decentralized stablecoins and collateral risk. *WBS Finance Group Research Paper Forthcoming* (2021).

[16]]makerDaoWhitepaper MakerDAO. [n. d.]. *The Maker Protocol: MakerDAO's Multi-Collateral Dai (MCD) System.* Technical Report. https://makerdao.com/en/whitepaper/#abstract

[17] Daniel Perez, Sam M Werner, Jiahua Xu, and Benjamin Livshits. 2021. Liquidations: DeFi on a Knife-edge. In *International Conference on Financial Cryptography and Data Security.* Springer, 457–476.

[18] Kaihua Qin, Liyi Zhou, Pablo Gamito, Philipp Jovanovic, and Arthur Gervais. 2021. An empirical study of defi liquidations: Incentives, risks, and instabilities. In *Proceedings of the 21st ACM Internet Measurement Conference.* 336–350.

[19] R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

[20] Robert Leshner, Geoffrey Hayes. 2019. *Compound: The Money Market Protocol.* Technical Report. https://compound.finance/documents/Compound.Whitepaper.pdf

[21] Patrick Royston and Mahesh KB Parmar. 2013. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology* 13, 1 (Dec. 2013). https://doi.org/10.1186/1471-2288-13-152

[22] Kevin Sekniqi, Daniel Laine, Stephen Buttolph, and Emin G¨un Sirer. 2020. Avalanche Platform. https://assets.website-files.com/5d80307810123f5ffbb34d6e/6008d7bbf8b10d1eb01e7e16_Avalanche%20Platform%20Whitepaper.pdf

[23] Xiao Shou, Tian Gao, Dharmashankar Subramanian, and Kristin P. Bennett. 2023. Multi-Label Event Prediction in Continuous Time. To appear.

[24] Xiao Shou, Georgios Mavroudeas, Alexander New, Kofi Arhin, Jason N. Kuruzovich, Malik Magdon-Ismail, and Kristin P. Bennett. 2019. Supervised Mixture Models for Population Health. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* 1057–1064. https://doi.org/10.1109/BIBM47256.2019.8983339

[25] Harmony Team. 2020. Harmony Technical Whitepaper. https://harmony.one/whitepaper.pdf

[26] Terry M. Therneau and Patricia M. Grambsch. 2000. *Modeling Survival Data: Extending the Cox Model.* Springer, New York.

[27] Terry M Therneau. 2022. *A Package for Survival Analysis in R.* https://CRAN.R-project.org/package=survival R package version 3.4-0.

[28] Mark Tyneway. 2020. Optimism. https://github.com/ethereum-optimism/optimism

[7] https://idea.rpi.edu/

[29] Abraham Wald. 1943. Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large. *Trans. Amer. Math. Soc.* 54, 3 (1943), 426–482. http://www.jstor.org/stable/1990256

[30] Jiahua Xu and Nikhil Vadgama. 2022. From banks to DeFi: the evolution of the lending market. In *Enabling the Internet of Value*. Springer, 53–66.

A TRANSACTION SUMMARIES BY MARKET

In section 2.1 we presented a table summarizing the transaction data that was used in this paper from the Aave Ethereum market. Since the data from other markets was also used for parts in this paper (e.g. section 3.3), we feel some might be interested in seeing the data summaries from other Aave markets. These are included below, with the market name, Aave version, and date-range of data included in the caption for each table.

Table 13. Summary of transaction types from Aave's **Avalanche V2** Market collected from March 31, 2021 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	318,227	$582,645.30	$1,200.00
Deposit	1,027,789	$383,204.90	$3,233.96
Redeem	783,107	$501,630.90	$5,660.20
Repay	199,300	$929,626.30	$5,650.52

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	10,900	$6,452.82	$6,873.20

Table 14. Summary of transaction types from Aave's **Polygon V2** Market collected from March 31, 2021 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	1,704,749	$29,383.27	$499.93
Deposit	13,338,917	$15,494.35	$7.22
Redeem	6,030,779	$14,106.76	$407.73
Repay	1,297,373	$31,252.15	$777.32

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	64,762	$7,296.67	$7,828.70

Table 15. Summary of transaction types from Aave's **Arbitrum V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	45,866	$2,572.83	$5.99
Deposit	119,543	$2,424.03	$11.07
Redeem	60,371	$3,838.99	$66.45
Repay	26,120	$3,966.26	$69.90

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	1,046	$1,502.62	$1,571.00

B COIN TYPES IN AAVE'S ETHEREUM MARKET

Table 16. Summary of transaction types from Aave's **Avalanche V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	74,243	$208,964.30	$711.14
Deposit	175,990	$129,160.30	$228.43
Redeem	85,420	$198,728.40	$1,759.38
Repay	50,476	$294,342.80	$1,792.30

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	1,999	$7,197.06	$7,618.99

Table 17. Summary of transaction types from Aave's **Fantom V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	11,591	$3,269.36	$200.02
Deposit	25,654	$4,598.34	$102.23
Redeem	17,794	$6,120.27	$213.73
Repay	9,580	$3,689.34	$299.99

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	475	$1,553.61	$1,655.37

Table 18. Summary of transaction types from Aave's **Harmony V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	4,064	$5,496.71	$199.98
Deposit	18,379	$3,302.44	$20.87
Redeem	7,791	$5,708.04	$104.40
Repay	3,602	$4,994.46	$86.14

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	269	$197.23	$210.21

Table 19. Summary of transaction types from Aave's **Optimism V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	88,759	$44,458.98	$1.00
Deposit	221,326	$23,822.92	$8.35
Redeem	97,036	$24,175.12	$23.14
Repay	40,934	$81,033.41	$19.43

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	824	$1,834.86	$1,961.56

Table 20. Summary of transaction types from Aave's **Polygon V3** Market collected from March 12, 2022 to October 01, 2022

Transaction Type	Occurrences	Mean Value (USD)	Median Value (USD)
Borrow	203,353	$27,890.93	$1,174.48
Deposit	287,995	$3,305.92	$50.00
Redeem	157,448	$5,062.66	$141.45
Repay	215,159	$26,196.11	$780.54

Transaction Type	Occurrences	Mean Principal (USD)	Mean Collateral (USD)
Liquidation	3,815	$1,302.76	$1,371.30

Coin Symbol	Coin Type
DAI	Stable
LINK	Non-Stable
Aave	Non-Stable
WBTC	Non-Stable
SNX	Non-Stable
USDC	Stable
TUSD	Stable
USDT	Stable
SUSD	Stable
BUSD	Stable
WETH	Non-Stable
YFI	Non-Stable
UNI	Non-Stable
BAT	Non-Stable
REN	Non-Stable
ENJ	Non-Stable
KNC	Non-Stable
MANA	Non-Stable
MKR	Non-Stable
ZRX	Non-Stable
CRV	Non-Stable
GUSD	Stable
BAL	Non-Stable
XSUSHI	Non-Stable
RENFIL	Non-Stable
RAI	Stable (Not pegged to USD)
AMPL	Non-Stable
PAX	Non-Stable
DPI	Non-Stable
FRAX	Stable
FEI	Stable
ENS	Non-Stable
UST	Non-Stable
CVX	Non-Stable
1INCH	Non-Stable
LUSD	Stable
STETH	Non-Stable