7-8-2023

Social Media Attention and the “Death” of Cryptocurrency: A Hazard Model Perspective

Peng Xie
California State University, East Bay, peng.xie@csueastbay.edu

Hongwei Du
California State University, East Bay, hongwei.du@csueastbay.edu

Social Media Attention and the "Death" of Cryptocurrency: A Hazard Model Perspective

Short Paper

Peng Xie
California State University, East Bay
25800 Carlos Bee Blvd, Hayward, CA 94542, United States
peng.xie@csueastbay.edu

Hongwei Du
California State University, East Bay
25800 Carlos Bee Blvd, Hayward, CA 94542, United States
hongwei.du@csueastbay.edu

Abstract

Include This paper studies the survival of cryptocurrencies and their association with the social media attention they receive. The death of a cryptocurrency is defined based on the discontinuation of trading activities and modeled using Kaplan – Meier Survivor Function and the Cox survival regressions. Using data collected from coinmarketcap.com and bitcointalk.org, we find that social media attention is a very relevant influencer for the death hazard. Specifically, the death hazard of a cryptocurrency is estimated to increase by 0.5% - 1% for each additional trading day without any social media mention. We also find that high-quality social media mentions are more effective in reducing the death hazard. The theoretical and practical implications of the findings are discussed in the paper.

Keywords: Social media, cryptocurrency, blockchain, survival analysis

Introduction

The development of blockchain technology introduced new ways to store information, and to make transactions in a decentralized manner. The creation of Ethereum in 2015 further fueled the growth of this field by enabling "smart contract" to be stored on the blockchain and interacted by the users. Various innovative services made possible by the smart contract engendered many different types of cryptocurrencies used to pay for these services.

Specialized exchanges are created to facilitate the trading of cryptocurrencies. One phenomenon occurred on these exchanges that we never see in a traditional stock exchange is the gradual cessation of trading of some cryptocurrencies that marks their "death". ThiIn this study, we model the "death" of cryptocurrencies using a time-dependent survival analysis and aim to investigate the factors that potentially lead to elevated hazard of death.

A relevant and comparable event in the stock market is delisting, which refers to the termination of trading of an asset in exchanges. Usually a stock may be delisted in mainly three ways: (1) delisted by SEC for rule violation; (2) voluntary delisting; and (3) delisted by exchanges due to unsatisfactory financial performance (Sanger and Peterson 1990). Delisting may also occur due to merger and acquisition, bankruptcy, liquidation, or migration to another exchange (Shumway and Warther 1999).

The stock market delisting events resemble cryptocurrency death in certain ways. First, it is found that the stocks to be delisted generate significantly negative returns compared to other stocks just before the delisting (Kashefi Pour and Lasfer 2013), and similar patterns can also be observed in the cryptocurrency market, where the capitalization quicky drop to near zero before death. Second, the liquidity of delisted stocks dramatically worsens. The investors may only use Pink Sheets OTC to trade the stocks in United States after delisting, and in some countries (such as United Kingdom), the stocks become private and completely illiquid after delisting (Leuz et al. 2008; Marosi and Massoud 2007). Similarly, dead cryptocurrencies can still be traded without using exchanges, but the liquidity is significantly worsened.

There are also noticeable differences between stock delisting and cryptocurrency death. First, the companies behind the listed stocks must comply with the SEC's listing standard, and a stock will be delisted if the company fails to meet these standards. However, currently no cryptocurrency asset is registered with SEC and there is no government regulation against unlawful or fraudulent offerings. Second, in the stock market, the delisting decision is publicly announced. In contrast, there is no "death announcement" in cryptocurrency market. Third, delisting is a terminal state in the stock market, at least in the short run. However, it is entirely possible that dead cryptocurrencies are resurrected after a while.

The focus of this study is to model the lifecycle and death of cryptocurrency using a survival model framework and explore the association between social media attention and death hazard. The social media framing of a cryptocurrency will impact the public perception of its collective salience and reflect the prevailing views of the community (Sheng and Lan 2019). Investors also trust mass media more than official accounts controlled by stakeholders because the mass media is more neutral and unbiased (Sheng and Lan 2019). We argue that frequent social media mentions indicate public attention, which implies investors' interest in cryptocurrency and the relevance of its functionality. In contrast, a lack of social media attention is an indicator of reduced interest and fading relevance, leading to high hazard of death.

The social media mention data is obtained from the Altcoin child message board in the leading cryptocurrency online community Bitcointalk.org. And the trading-related data is obtained from coinmarketcap.com, and it is used to determine the death time of cryptocurrencies and to compute the trading-related control variables. Then a time-dependent covariate hazard model is estimated to explore the predictors of the hazard.

To preview the results, we found that the death hazard of a cryptocurrency is estimated to increase by 0.5% - 1% for each additional trading day without any social media mention. We also find that that high-quality social media mentions (i.e., endorsed by many social media users) will further reduce the death hazard.

This paper offers useful practical insights. Public attention to cryptocurrency is associated with its continued trading. To stimulate trading, the development teams could intermittently release updates or announcements on social media to prevent trading halt.

Data

The data regarding the trading dynamics and the lifecycle of cryptocurrencies is downloaded from coinmarketcap.com. We collected daily price, trading volume and market capitalization (all measured in terms of US dollars) for 4,529 different cryptocurrencies enlisted by coinmarketcap.com before the end of the data collection period (May. 31, 2021). In this sample, the earliest enlisting time is Dec. 27, 2013, which is the date Bitcoin is incorporated by this coinmarketcap.com. It has been indicated in the literature that coinmarketcap.com is a reliable data source for cryptocurrency trading data (Vidal-Tomás 2022).

According to the classical survival analysis framework, a cryptocurrency either survived beyond the data collection period (i.e., censored by the dataset) or "died" before it. In this research, the death of a cryptocurrency is defined based on the cessation of trading (i.e., trading volume gradually declines to zero). A cryptocurrency is treated as "dead" on day t_0 if the trading volume is zero across t_0, t_1, and t_2.

Among the 4,529 cryptocurrencies in the sample, some are never traded since enlisted, and therefore are marked "death" from the first day, and some reaches its death too soon although actively traded before death. These short-lived cryptocurrency samples are excluded from the analysis due to (1) the death could be potentially ascribed to the ICO process (i.e., the failure of the project launching) but not the factors of interest in this research; and (2) many of the covariates used to estimate the death hazard is calculated using moving average and they cannot be calculated for short-lived cryptocurrencies. In this research, only 1,980 cryptocurrencies surviving at least 60 days are retained in the sample.

To obtain the social media attention received by each cryptocurrency, we downloaded the social media discussions posted Bitcointalk.org, which is one of the most popular online communities for cryptocurrency investors. It has been used in many cryptocurrency related studies in the literature (Mai et al. 2018; Xie et al. 2020). On this online community, there is a child message board dedicated to the social media discussions for “altcoins” (i.e., cryptocurrencies other than Bitcoin).

For each of the cryptocurrencies in the sample, we identified their social media mentions by searching for their ticker symbols in the social media messages. From the social media mentions, we calculated the following variables to capture the public attention received by each cryptocurrency: (1) $ifAnyMention_t$, which is binary variable set to 1 if there is at least one mention during the past 30 days; (2) $nMentions_t$, which is the number of mentions received during the past 30 days; (3) $Sentiment_t$, which is the percentage of bearish words in all messages during the past 30 days that mentions a particular cryptocurrency (Tetlock 2007; Tetlock et al. 2008; Xie et al. 2020); (4) $nDaysSinceLastMention_t$, which is the number of days since a particular cryptocurrency is last mentioned; and (5) $nComenions_t$, which is the number of times a particular cryptocurrency is co-mentioned (i.e., mentioned simultaneously) with other cryptocurrencies in the same message.

Empirical Analysis

Survivor Function Estimation using Kaplan – Meier Estimator

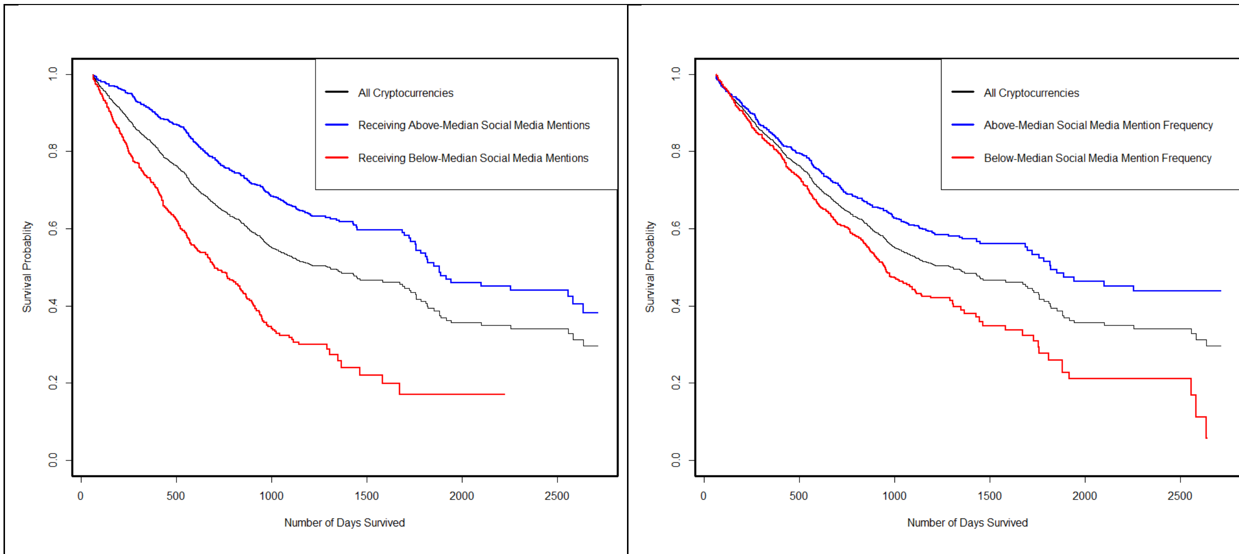
Each type of submission (completed research This section describes a model-free evidence using the Keplan – Meier estimator (KM estimator), which is a widely used method in survival analysis to estimate the survivor function (Allison 2010; Calabuig et al. 2021). We observe each of the n cryptocurrencies in the sample ($i = 1 \dots n$) every day discretely (at times $t = 1,2,3, \dots, T_i$) starting at $t = 1$ (the first day it is listed on coinmarketcap.com) and ending at $t = T_i$ (the last observation date for cryptocurrency i). Data beyond T_i (i.e., $t > T_i$) are no longer unobservable for cryptocurrency i due to the end of the data collection period. Following the standard survival analysis literature (Rao et al. 1998), if a cryptocurrency i’s death event occurred before or on T_i, then we define t_d as its death time. If a cryptocurrency’s death event has not occurred at $t = T_i$, then it is censored.

Each cryptocurrency in our sample is listed on coinmarketcap.com on a different date, but they are either died or censored no later than the end of the data collection period, May 31, 2021. Following Allison (Allison 2010), we define the survivor function to be $S(t) = \Pr(t_d > t)$, which is the probability that a cryptocurrency survives beyond t. KM estimator is a commonly used estimation for the survivor function, and it is defined as:

$$\hat{S}(t) = \prod_{t_1<t_j<t_k} (1 - \frac{d_j}{n_j}), \tag{1}$$

where $t_1 < t_2 < \cdots < t_k$ are k distinct time point when at least one death event occurred. At each death time t_j, d_j is the number of death events occurred at that time point, and n_j is the number of cryptocurrencies at risk of death at that time point (i.e., the number of cryptocurrencies that have neither died nor censored before time t_j).

To visually illustrate how social media mentions of a cryptocurrency might predict its chance of survival, we compare (1) the survivor function estimation between the cryptocurrencies receiving above-median social media mentions and these receiving below-median social media mentions; and (2) the survivor function estimation between cryptocurrencies with above-median social median mention frequency (i.e., the average number of social media mentions received per day) and these with below-median social media mention frequency. The comparison between these estimated survivor functions is presented in Figure. 1.



(a) (b)

Figure 1. Kaplan – Meier Survivor Function Estimation Comparison

It is evident from Figure. 1 that cryptocurrencies with above-median social media mentions and mention frequency always enjoy higher survival probability than others. This model-free evidence supports our argument that social media mentions positively predicts the longer survival. However, in this analysis, we only utilized the aggregated information (i.e., total number of received mentions and mention frequency) in distinguishing between long-surviving and short-surviving cryptocurrencies. In the next section, we will estimate a time-dependent covariate hazard model to gain deeper insights.

COX Hazard Model Estimation

For all cryptocurrencies in our sample, we keep observing them each day until (1) a death event occurred; or (2) the observation is censored. We then define $Y_i = \min(t_d, T_i)$ to be the last observation date for cryptocurrency i.

Also, the status of cryptocurrency i at the last observation date N_i is set to 1 if $Y_i = t_d$ (a death event occurred), and 0 if $Y_i = T_i$ (the cryptocurrency survived beyond the censoring date). We also define the discrete time hazard rate for cryptocurrency i at time t to be $P_i^t = \Pr(t_d = t | X_i^1, \dots, X_i^{Y_i}, t_d \geq t)$, which is the probability that cryptocurrency i died at time point t conditional on its time-dependent covariates and the fact the it survived all previous time points before t.

The likelihood L for cryptocurrency i's observation is then shown as follows:

$$L(N_i | X_i^1, \dots, X_i^{Y_i}) = \begin{cases} P_i^{Y_i} \prod_{t=1}^{Y_i - 1} (1 - P_i^t), & \text{if } N_i = 1 \\ \prod_{t=1}^{Y_i} (1 - P_i^t), & \text{if } N_i = 0 \end{cases}$$

$$= [P_i^{Y_i} \prod_{t=1}^{Y_i - 1} (1 - P_i^t)]^{N_i} [\prod_{t=1}^{Y_i} (1 - P_i^t)]^{1 - N_i} \tag{2}$$

And the log-likelihood function can be written as below after rearrangement:

$$L(N_i | X_i^1, \dots, X_i^{Y_i}) = N_i \log\left(\frac{P_i^{Y_i}}{1 - P_i^{Y_i}}\right) + \sum_{t=1}^{Y_i} \log(1 - P_i^t) \tag{3}$$

We also follow the standard assumption that the death events across different cryptocurrencies are independent, which then gives the complete log-likelihood function for the entire sample:

$$L\left(N_1, \dots, N_n \middle| X_1^1, \dots, X_1^{Y_1}, \dots, X_n^1, \dots, X_n^{Y_n}\right) = \sum_{i=1}^{n} \log L\left(N_i \middle| X_i^1, \dots, X_i^{Y_i}\right) \quad (4)$$

Finally, the hazard rate is parameterized for model estimation using the following logit regression equation:

$$\log\left(\frac{P_i^t}{1-P_i^t}\right) = \boldsymbol{\beta}^T \boldsymbol{X}_i^t => P_i^t = \frac{1}{1+\exp\left(-\beta^T X_i^t\right)} \quad (5)$$

Following the survival analysis literature, $\frac{P_i^t}{1-P_i^t}$ is the hazard of a cryptocurrency dies at time point t and the contribution to the hazard from each time-dependent covariate is $\exp(\beta)$. A $\exp(\beta)$ larger than 1 indicates a particular covariate is positive related to the hazard of dying, and vice versa.

In this analysis, we aim to investigate if social media mentions cryptocurrency predict its survival. A series of social media related variables are created and they include: (1) $AnyMention_t$, which is a binary variable set to 1 if there is at least one social media mention during the past 30 days; (2) $NumMention_t$, which is the logarithm of the count of mentions during the past 30 days (logarithm operation is applied to smooth the distribution of the variable); (3) $NegSentiment_t$, which is the total number of bearish words divided by the total number of words in all messages mentioning a particular cryptocurrency during the past 30 days (this variable is set to 0 if $AnyMention_t = 0$); (4) $DaysSinceLastMention_t$, which is the number of days since the cryptocurrency is last mentioned in some social media messages (this variable is set to 0 if $AnyMention_t = 0$); (5) $Merit_t$, which is the total number of merits received by all messages posted during the past 30 days. The merit score is the endorsement for the posted messages. If a message is perceived to be highly informative and accurate, it tends to receive more merit scores.

Please note that the effect of zero social media mention during the past 30 days is captured by the coefficient of the variable $AnyMention_t$ since other related variables are set to zero. Also note that besides $DaysSinceLastMention_t$, all the other social media related variables are calculated using monthly moving average. We also include various market dynamics variables to control the influence of cryptocurrency trading on its survival. These market dynamics variables include (1) $Return_t$, which is the cumulative returns over the past 30 days; (2) $MarketCap_t$ and $TradingVolume_t$, which are the logarithm of the 30-day moving average market capitalization and trading volume; (3) $VolumeStdev_t$, which is the standard deviation of the trading volume over the past 30 days; (4) $NumHighReturn_t$, and $NumLowReturn_t$, which are the number of days in the past 30 days that experiences a return larger than 10% and the number of days in the past 30 days that experiences a return lower than -10% (since cryptocurrency market is volatile, large returns is common especially for small-cap cryptocurrencies); (5) $Volatility_t$, which is sum of squared returns over the past 30 days; and (6) $ReturnSlope_t$, $CapSlope_t$, $VolumeSlope_t$, and $VolatilitySlope_t$, which are the slope obtained by regressing $Return_t$, $MarketCap_t$, $TradingVolume_t$, and $Volatility_t$ on t using the past 30 days' data, respectively. These slope variables capture the general short-term trend of these variables.

Please note that all variables mentioned above except $DaysSinceLastMention_t$ are calculated using the past 30-day moving average, so that this model captures the short-term influence of the variables on the death hazard. We also explore the long-term influence by estimating a similar model where all variables are calculated using cumulative data (i.e., variables are calculated using a time window spanning from $t = 1$ to the current day). For example, $AnyMention_t$ is coded as 1 if there is any social media mention since it is first listed on coinmarketcap.com until time point t, and 0 otherwise. Note that the variables $NumHighReturn_t$ and $NumLowReturn_t$ are replaced by $PctHighReturn_t$ and $PctLowReturn_t$ respectively in the long-term model, which measure the percentage of trading days with a return greater (lower) than 10% (-10%). The summary statistics of the variables are presented in Table 1 (logarithm is not applied in this table).

	30-day moving average					Cumulative				
	Max	Min	Median	Mean	Std.Dev	Max	Min	Median	Mean	Std.Dev
DaysSinceLastM	2,683	0	61	184.769	282.156	2,683	0	61	184.769	282.156

AnyMention	1	0	0	0.291	0.454	1	0	1	0.882	0.323
NumMention	7,061	0	0	9.933	115.424	100,363	0	16	355.514	3374.023
NegSentiment	0.2	0	0	0.002	0.007	0.094	0	0.006	0.007	0.006
Return	165.652	-0.211	0.005	0.028	1.042	30.165	-0.140	0.007	0.022	0.242
MarketCap	1.09e+12	0	3,842,913	4.59e+08	1.13e+10	1.23e+11	193.632	6,878,295	1.76e+08	2.11e+09
TradingVolume	1.58e+11	1	113,070.3	9.32e+07	1.87e+09	1.85e+10	5.166667	263,046.8	2.17e+07	2.96e+08
VolumeStdev	5.55e+10	0	60,421.94	3.17e+07	5.93e+08	3.44e+10	5.505208	421,157.8	3.53e+07	5.03e+08
Volatility	2.46e+07	0	0.308	971.621	127,651.9	1098.379	0	0.199	0.500	9.672855
NumHighReturn	19	0	3	3.610	2.903					
NumLowReturn	22	0	3	3.488	2.924					
PctHighReturn						0.537	0	0.115	0.126	0.065
PctLowReturn						0.710	0	0.116	0.127	0.070
ReturnSlope	31.977	-31.908	0	0.00008	0.119	0.955	-6.526	0	-0.0003	0.021
CapSlope	0.377	-0.487	-0.009	-0.005	0.072	0.155	-0.109	-0.002	-0.002	0.015
VolumeSlope	0.233	-0.428	-0.004	-0.003	0.047	0.107	-0.171	-0.001	-0.001	0.010
VolatilitySlope	3180.349	-3180.331	-0.00004	0.011	10.247	140.401	-101.036	-0.0001	-0.002	0.421

Table 1. Descriptive Statistics

The estimation results are presented in Table 2. The short-term effect model using variables calculated by a 30-day moving average window are shown in columns (1) and (2) and the long-term effect model using variables calculated by a cumulative time window is shown in columns (3) and (4). The result provides support for our argument that social media mentions could influence the hazard of cryptocurrency death. Specifically, the coefficient estimates for $DaysSinceLastMention_t$ in both models range from 0.0005 to 0.001 and are statistically significant, meaning that on average, one more day without social media mention is associated with a 0.5% to 1% increase in the death hazard. The hazard of $AnyMention$ is not statistically significant in the short-term model, meaning that the short-term social media attention loss does not significantly increase the death hazard. However, the hazard of $AnyMention$ in the long-term effect model is statistically significant at 0.614, meaning that compared to cryptocurrencies not mentioned in social media at all in the entire history, the death hazard of cryptocurrencies receiving at least one mention is 38.6% lower (38.6% = (1 - 0.614) × 100%) on average. We also found that the coefficients of the message merit are consistently negative and statistically significant at least at the 5% level. This additional result indicate that higher quality social media mentions will further reduce the death hazard. These additional results are added to the paper.

	30-day moving average (short-term influence)		**Cumulative (long-term influence)**	
	Coefficient Estimate	Hazard	Coefficient Estimate	Hazard
	(1)	(2)	(3)	(4)

DaysSinceLastMention	0.0007*** (5.062)	1.001***	-0.001*** (8.404)	1.001
AnyMention	-0.993 (-1.152)	0.370	-0.627 (-1.490)	0.534
NumMention	0.014 (0.085)	1.014	0.015 (0.358)	1.015
NegSentiment	-15.080 (-1.017)	2.828e-07	-3.047 (-0.460)	0.047
Return	0.0003 (1.168)	1.000	0.112 (0.451)	1.118
MarketCap	-0.004 (-1.686)	0.962	-0.120*** (-3.541)	0.887***
TradingVolume	-0.730*** (-9.715)	0.482***	-0.226** (-2.663)	0.798**
VolumeStdev	0.401*** (5.107)	1.493***	0.063 (0.781)	1.065
Volatility	-1.471e-05 (-1.712)	1.000	-0.007 (-0.403)	0.993
NumHighReturn	0.0408 (1.710)	1.042		
NumLowReturn	0.0472* (2.065)	1.048*		
PctHighReturn			-1.782 (-0.965)	0.168
PctLowReturn			7.236*** (4.181)	1,388
ReturnSlope	3.627*** (3.710)	37.590***	13.130 (0.867)	5.025e+05
CapSlope	-0.231 (-0.284)	0.794	-19.020* (-2.371)	5.512e-09
VolumeSlope	-11.230*** (-9.204)	1.323e-05***	-26.360* (-2.303)	3.555e-12
VolatilitySlope	0.0130 (0.650)	1.013	0.324 (0.443)	1.383
Merit	-0.249* (-2.447)	0.780*	-0.120** (-2.606)	0.887**
AuthorActivity	0.294 (1.558)	1.341	0.100 (1.113)	1.106
# Obs.	1,236,316		1,236,316	
Likelihood Ratio Test	Likelihood ratio = 1219; df = 17; p=0		Likelihood ratio = 705; df = 17; p=0	

Table 2. Hazard Model Estimation

The two models also show that as long as the cryptocurrency is mentioned, the count of the mentions and their sentiment does not significantly affect the death hazard as the coefficient estimates for $numMension$ and $negSentiment$ are not statistically significant. Note that the hazard is used to interpret the effect of binary variable such as $AnyMention$ since the changes in the hazard can be exactly calculated, while coefficient estimate is used to interpret all other non-binary discrete variables such as $DaysSinceLastMention_t$ due to the non-linear nature of the hazard function.

Predicting the Death of Cryptocurrency

To gauge the out-of-sample hazard model performance, we perform a prediction test following the same procedure used in predicting bankruptcy in the finance literature (Chava and Jarrow 2004). We use all cryptocurrencies listed before September 03, 2019, as the training set to train the survival model and all cryptocurrencies listed after September 03, 2019, as the testing set. There are 139 death events occurred in the test set and 553 death events occurred in the training set. Since there are a total of 692 death events in our sample of 1980 cryptocurrencies. The train/set division set aside roughly 20% of the death events in the test set and 80% of the death events in the training set.

After the model is trained on the training set, it is applied to the test set to calculate the $hazard = \exp(-\beta^T X_i^t)$ of each cryptocurrency across each day. Then during each day with at least one death event, we rank the cryptocurrencies at risk into deciles based on their hazard. Under this construction, the cryptocurrencies that are predicted to have high death hazards are placed into the first few deciles and the cryptocurrencies that are predicted to have low death hazard are placed into the last few deciles. The number of cryptocurrencies in each decile that are actually experiencing a death event in that day is counted and aggregate through all days covered in the test set.

The predicted result is presented in Table 3. Four deciles are used in Panel A and ten deciles are used in panel B. It is evident that most of the cryptocurrencies that experienced a death event have a high hazard with respect to others in that day. The model estimated with both long-term and short-term covariates and the short-term covariates only model performs relatively better than the long-term covariates only model.

We further test the prediction accuracy using the area under the ROC curve. Specifically, in each day with at least one death event, we predict the death of each cryptocurrency and generate the ROC curve for that day, then we calculate the AUROC (area under the ROC curve) to summarize the prediction accuracy. In all models, the mean AUROC across all death event days are above 80%, indicating a decent prediction performance. Also, the short-term covariate model outperforms the other models with a mean AUROC of 0.868.

Decile	**Model with Both Long-term and Short-term Covariates**	**Model with Short-term Covariates**	**Model with Long-term Covariates**
Decile	Percentage of the 139 death events in the test set in each decile		
Panel A: Using Four Deciles			
1	79.1%	79.9%	72.7%
2	13.7%	12.9%	18.0%
3	3.6%	5.0%	6.5%
4	3.6%	2.2%	2.9%
Panel B: Using Ten Deciles			
1	51.8%	54.7%	36.0%
2	19.4%	20.1%	25.9%
3	12.9%	7.2%	15.1%
4	3.6%	7.9%	9.4%
5	5.0%	2.9%	4.3%
6	0.7%	3.6%	2.9%
7	2.2%	1.4%	3.6%
8	1.4%	1.4%	2.2%
9	2.2%	0.7%	0%
10	0.7%	0%	0.7%

Table 3. Cryptocurrency Death Prediction Accuracy

Conclusion

This research demonstrates the viability of using survival models such as Kaplan – Meier Estimator and COX model to model the lifecycle of tradable assets in the cryptocurrency context. In the future, we plan to follow the route to explore other modeling methods. For example, because trading related data can be viewed as a sequential time series data, a recursive neural network could be used to capture the interdependency among events occurring in different time and attempt to achieve high "death" prediction accuracy.

References

Allison, P. D. 2010. "Survival Analysis Using SAS: A Practical Guide, Second Edition," Technometrics.

Calabuig, J. M., García-Raffi, L. M., García-Valiente, A., and Sánchez-Pérez, E. A. 2021. "Kaplan-Meier Type Survival Curves for COVID-19: A Health Data Based Decision-Making Tool," Frontiers in Public Health. (https://doi.org/10.3389/fpubh.2021.646863).

Chava, S., and Jarrow, R. A. 2004. "Bankruptcy Prediction with Industry Effects*," Review of Finance. (https://doi.org/10.1093/rof/8.4.537).

Kashefi Pour, E., and Lasfer, M. 2013. "Why Do Companies Delist Voluntarily from the Stock Market?," Journal of Banking and Finance. (https://doi.org/10.1016/j.jbankfin.2013.08.022).

Leuz, C., Triantis, A., and Yue Wang, T. 2008. "Why Do Firms Go Dark? Causes and Economic Consequences of Voluntary SEC Deregistrations," Journal of Accounting and Economics. (https://doi.org/10.1016/j.jacceco.2008.01.001).

Mai, F., Shan, Z., Bai, Q., Wang, X. (Shane), and Chiang, R. H. L. 2018. "How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis," Journal of Management Information Systems (35:1), p. 2018. (https://doi.org/10.1080/07421222.2018.1440774).

Marosi, A., and Massoud, N. 2007. "Why Do Firms Go Dark?," Journal of Financial and Quantitative Analysis. (https://doi.org/10.1017/s0022109000003331).

Rao, M. B., Klein, J. P., and Moeschberger, M. L. 1998. "Survival Analysis Techniques for Censored and Truncated Data," Technometrics. (https://doi.org/10.2307/1270658).

Sanger, G. C., and Peterson, J. D. 1990. "An Empirical Analysis of Common Stock Delistings," The Journal of Financial and Quantitative Analysis. (https://doi.org/10.2307/2330828).

Sheng, J., and Lan, H. 2019. "Business Failure and Mass Media: An Analysis of Media Exposure in the Context of Delisting Event," Journal of Business Research. (https://doi.org/10.1016/j.jbusres.2018.01.055).

Shumway, T., and Warther, V. A. 1999. "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect," Journal of Finance. (https://doi.org/10.1111/0022-1082.00192).

Tetlock, P. C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," Journal of Finance (62:3), pp. 1139–1168. (https://doi.org/10.1111/j.1540-6261.2007.01232.x).

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. 2008. "More Than Words: Quantifying Language to Measure Firms' Fundamentals," Journal of Finance. (https://doi.org/10.1111/j.1540-6261.2008.01362.x).

Vidal-Tomás, D. 2022. "Which Cryptocurrency Data Sources Should Scholars Use?," International Review of Financial Analysis. (https://doi.org/10.1016/j.irfa.2022.102061).

Xie, P., Chen, H., and Hu, Y. J. 2020. "Signal or Noise in Social Media Discussions: The Role of Network Cohesion in Predicting the Bitcoin Market," Journal of Management Information Systems (37:4), pp. 933–956. (https://doi.org/10.1080/07421222.2020.1831762).