Article

Cryptocurrencies and Tokens Lifetime Analysis from 2009 to 2021

Paul Gatabazi [1,2,*], Gaëtan Kabera [3], Jules Clement Mba [4], Edson Pindza [5,6] and Sileshi Fanta Melesse [7]

1   Department of Mathematics and Applied Mathematics, Faculty of Sciences, University of Johannesburg, Johannesburg 2006, South Africa
2   Department of Applied Statistics, School of Economics, College of Business and Economics, University of Rwanda, Huye 124, Rwanda
3   Department of Statistics, University of South Africa—UNISA, Pretoria 0003, South Africa; kaberg@unisa.ac.za
4   School of Economics, College of Business and Economics, University of Johannesburg, Johannesburg 2006, South Africa; jmba@uj.ac.za
5   Department of Mathematics and Statistics, Tshwane University of Technology, Pretoria 0001, South Africa; pindzaedson@yahoo.fr
6   Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria 0002, South Africa
7   School of Mathematics, Statistics and Computer Science, University of KwaZulu Natal, Pietermaritzburg 3209, South Africa; melesse@ukzn.ac.za
*   Correspondence: pgatabazi@uj.ac.za

Abstract: The success of Bitcoin has spurred emergence of countless alternative coins with some of them shutting down only few weeks after their inception, thus disappearing with millions of dollars collected from enthusiast investors through initial coin offering (ICO) process. This has led investors from the general population to the institutional ones, to become skeptical in venturing in the cryptocurrency market, adding to its highly volatile characteristic. It is then of vital interest to investigate the life span of available coins and tokens, and to evaluate their level of survivability. This will make investors more knowledgeable and hence build their confidence in hazarding in the cryptocurrency market. Survival analysis approach is well suited to provide the needed information. In this study, we discuss the survival outcomes of coins and tokens from the first release of a cryptocurrency in 2009. Non-parametric methods of time-to-event analysis namely Aalen Additive Hazards Model (AAHM) trough counting and martingale processes, Cox Proportional Hazard Model (CPHM) are based on six covariates of interest. Proportional hazards assumption (PHA) is checked by assessing the Kaplan-Meier estimates of survival functions at the levels of each covariate. The results in different regression models display significant and non-significant covariates, relative risks and standard errors. Among the results, it was found that cryptocurrencies under standalone blockchain were at a relatively higher risk of collapsing. It was also found that the 2013–2017 cryptocurrencies release was at a high risk as compared to 2009–2013 release and that cryptocurrencies for which headquarters are known had the relatively better survival outcomes. This provides clear indicators to watch out for while selecting the coins or tokens in which to invest.

Keywords: cryptocurrency; blockchain; survival function; risk; weight; hazard ratio

1. Introduction

Cryptocurrencies are digital currencies in which transactions are verified and records maintained by decentralized systems known as blockchains. Blockchains use cryptography or a third-free peer-to-peer electronic system, rather than a centralized trade in which transactions are made by the banks (Angela 2016). Transacting a cryptocurrency is mobile, non-taxable and does not require bank intermediary. Blundell-Wignall (2014) points that digital cash cannot have multiple copies. Hence, a cryptocurrency cannot be used more than once, unlike the bank services where multiple transfers are common.

Cryptocurrecies operate either as coins or tokens. Crypto-coins are native from their own blockchain whilst tokens are built on top of another existing blockchain (Wu et al. 2018). An important number of cryptocurrencies use Euthereum network as an alternative to standalone. Many other known networks include Waves, Stellar, Nem, Couterparty, Bitshares, Achain, Omni, Neo, Ardor, Qtum, Icon and Ubiq. The discretion in cryptocurrency transaction makes a more secure and reliable modern mode of payment as suggest manuscripts such as Blundell-Wignall (2014), Hendrickson et al. (2016) or Urquhart (2016).

The cryptocurrencies release started with the Bitcoin in 2009 Blundell-Wignall (2014). The other popular cryptocurrencies that followed include Litecoin (2011), Peercoin (2012), Ripple (2012), Alphacoin (2013–2014) and Aircoin (2014–2016). As for the year 2021, the new cryptocurrencies include Bogged Finance (2021) and Recharge Finance (2020). The total number of cryptocurrencies in 2021 exceeds 6000 in which around 2000 are inactive (or dead) (Best 2022). The high rate of death of cryptocurrencies has been a barrier of many investors. The risk of collapsing of a number of cryptocurrencies has persuaded some governments to ban such digital financial markets with argument that cryptocurrency trade could facilitate illegal transactions and disrupting the government activities (Hendrickson et al. 2016). However, El Salvador officially adopted Bitcoin as legal tender on 9 June 2021, making it the first country to do so (Thul 2021). In some other countries such as Panama, Brasil, Paraguay, Mexico and Argentina, the governments are hand in hand with the researchers on sustainability of Bitcoin for a future adoption (Thul 2021).

In 2020, the globe was under a relatively rising situation of the outbreak of COVID-19 and at the same time, the relatively high degree spillover of Bitcoin was observed (Xiao et al. 2021). The trade of cryptocurrencies was then globally affected (Youssef et al. 2021). Azimli (2020) suggests that COVID-19 pandemic impacts the financial market through the high level of economic policy uncertainty. The unknown future situation of COVID-19 leads to low cash flow expectations, resulting in possible stock market depreciation (Azimli 2020).

Several studies on cryptocurrencies emerged in many domains such as Mathematical Sciences, Financial Economics and Engineering. Mathematical and statistical modeling on cryptocurrency is found for example in (Gatabazi et al. 2019a, 2019b, 2019c, 2019d) and Chan et al. (2017). Mba and Wang (2019) used copula in describing cryptocurrencies in financial economics framework, a field in which many other manuscripts including Urquhart (2016) and Angela (2016) analyse cryptocurrencies. However, the scientific analysis on the lifetime of cryptocurrencies is not yet popular. To understand the event history analysis of cryptocurrencies and the associated factors may adopt another insight by which the researcher may predict accurately the future of these digital currencies.

The present study uses the time-to-event data analysis for estimating the life-time and the factors of death for cryptocurrencies during the study time ranging from 2009 to 2021.

In addition to the introductory section, the study is subdivided as follows: Section 2 presents the methodology of the study. Section 3 presents the main results and interpretation and Section 4 gives a conclusion.

2. Methodology

This part introduces the time-to-event analysis, discusses the time-to-event regression analysis and then presents the dataset of interest. Preliminary analysis will be displayed by the graphs and the useful statistical tests.

2.1. Concept of the Time-to-Event Data Analysis

The time-to-event analysis or survival analysis aims at making the inference of the time elapsed between the onset of observations until the occurrence of some event of interest. The related regression model expresses the dependence of time-to-event on predictor variables. Methods used in general statistical analysis, in particular in regression analysis, are not directly applicable to survival data due to censoring and truncation. Hosmer et al. (2008) describe three types of censoring: *right censoring* arising when an individual is not subject to the event until the end of study due to either loss to follow up, or the event

has not occurred at the end of the study, or the event has occurred from another cause not related to the cause of interest. *Left censoring* arises when an individual experienced an event before the onset of the study. *Interval censoring* refers to when the event occurs within some interval in the study time, or the individual dropped out or observed the event at an unknown time before study termination for reasons unrelated to the study, or the individual was lost to follow-up in an interval between two specified time points. Two types of truncation described in Klein and Moeschberger (2003) are the *left truncation* and the *right truncation*. Left truncation occurs when subjects under a survival study have been at risk before the study time. Right truncation arises when interest is only on individuals who have experienced the event by a specified future time before study termination. In this study, interest will be only on right censoring.

In time-to-event analysis, a non-negative random variable representing the time-to-event is generally characterized by three fundamental functions: the probability density function (for continuous random variables) or probability mass function (for discrete random variables), the survival function and the hazard function as detailed in Hosmer et al. (2008). The hazard function is also known as risk function or intensity rate (Hosmer et al. 2008). Any of these three functions can be uniquely determined from at least one of the other two functions (Collet 2003; Hosmer et al. 2008).

2.2. Comparison of Two or More Groups of Survival Data

Two or more groups survival time may be compared by using the plots of the survival functions in one system of axes. Log-rank and Wilcoxon tests are popular tests for comparing survival functions (Collet 2003; Klein and Moeschberger 2003). The tests are based on the following hypotheses:

H0: *No difference in survival experiences of the individuals in the groups,*

H1: *There is difference in survival experiences of the individuals in the groups.*

In several studies, statistical significance is based on comparing p-values to a specified level of significance, α Generally $\alpha = 0.05$ or $\alpha = 0.01$. In this study, we prefer using the interpretation described in Collet (2003) that is summarised in Table 1. The same interpretation was used in Gatabazi and Kabera (2015), and Gatabazi et al. (2018, 2019e, 2020a).

Table 1. Evidence for or against H_0 based on comparing the p-value with the level of significance $\alpha = 0.05$.

p-Value (P)	Interpretation
$P > 0.1$	No evidence to reject the null hypothesis
$0.05 < P \leq 0.1$	Slight evidence against the null hypothesis
$0.01 < P \leq 0.05$	Moderate evidence against the null hypothesis
$0.001 \leq P \leq 0.01$	Strong evidence against the null hypothesis
$P < 0.001$	Overwhelming evidence against the null hypothesis

The log-rank test is better if proportional hazards can be assumed (Collet 2003). In such situation, the plots of survival functions do not cross one another. The Wilcoxon test is suitable when there is no proportional hazards assumption. Here, the survival curves of some groups cross one another (Collet 2003).

2.3. Cox Proportional Hazards Model (CPHM)

Assume p fixed covariates with values $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ for $i = 1, \ldots, n$ where n is the number of observations.

The CPHM is given by

$$h(t \mid \mathbf{x}_i) = h_0(t) \exp(\boldsymbol{\beta}' \mathbf{x}_i) \tag{1}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$ is a p-dimensional vector of model parameters and and $h_0(t)$ is the baseline hazard function, that the hazard function when all the covariates are set to zero. The quantity

$$\psi_k = e^{\beta_k} \tag{2}$$

is called "*hazard ratio*", and is reported in applied studies as it is easier to interpret than the model parameters or log-hazard ratio $\beta_k = \ln \psi_k$ for $k = 1, \ldots, p$ (Collet 2003).

Parameter estimation for model (1) with no tied events is conducted using partial likelihood introduced by Cox (1972).

Three approaches of approximating the partial likelihood in presence of tied event are suggested by Breslow (1974), Efron (1977) and Cox (1972). In practice, the three approximations of the partial likelihood function lead to similar results (Collet 2003). Many statistical packages, including STATA, provide options for using each of the above approximations. In STATA, Beslow is taken as the default.

2.4. Aalen Additive Hazards Model (AAHM)

The AAHM at time t of the ith of n individuals is given by

$$h[t \mid \mathbf{x}_i(t)] = \beta_0(t) + \beta_1(t)x_{i1}(t) + \beta_2(t)x_{i2}(t) + \ldots + \beta_p(t)x_{ip}(t) \tag{3}$$

where $\boldsymbol{\beta}(t) = (\beta_0(t), \beta_1(t), \ldots, \beta_p(t))'$ is the vector of parameter functions that may be estimated and $\mathbf{x}_i(t) = (x_{i1}(t), x_{i2}(t), \ldots, x_{ip}(t))'$ is the vector of covariates. The parameter function $\beta_0(t)$ is the baseline hazard function, that is the hazard when all the covariate functions are set to zero.

Aalen et al. (2008, p. 157) argue that, for computation stability, estimation in model (3) should be based on the cumulative parameter functions

$$B_k(t) = \int_0^t \beta_k(v)dv, \tag{4}$$

$k = 0, 1, 2, \ldots, p$. Clearly, if $\beta_k(t)$ is constant, say $\beta_k(t) = \beta_k$, then

$$B_k(t) = \int_0^t \beta_k dv = \beta_k t \quad \text{which is represented by a straight line.}$$

Proposition 1. *Let*

$$Y_i(t) = \begin{cases} 1, & \text{if individual } i \text{ is at risk at time } t \\ 0, & \text{otherwise.} \end{cases}$$

Assume that $dN_i(t)$ and $dM_i(t)$ are respectively a response variable and a random error terms of the ith individual. Model (3) leads to the form

$$dN_i(t) = \sum_{k=0}^p Y_i(t)x_{ik}(t)dB_k(t) + dM_i(t) \tag{5}$$

where $x_{i0} = 1$.

The proof of Proposition 1 can be found for example in Gatabazi et al. (2020b).

Thus, Model (5) has the form of a multiple linear regression model for the ith individual with covariates $Y_i(t)x_{ik}$ and parameters $dB_k(t)$ for $k = 0, 1, 2, \ldots, p$ and $i = 0, 1, 2, \ldots, n$.

Model (5) can be written in matrix form as

$$d\mathbf{N}(t) = \mathbf{X}(t)d\mathbf{B}(t) + d\mathbf{M}(t) \tag{6}$$

where

$d\mathbf{N}(t)$ is the $n \times 1$ vector of observations $dN_i(t)$

$\mathbf{X}(t)$ is the $n \times (p+1)$ design matrix with ith row given by $Y_i(t), Y_i(t)x_{i1}(t), \ldots, Y_i(t)x_{ip}(t)$

$d\mathbf{B}(t) = \left(dB_0(t), dB_1(t), \ldots, dB_p(t)\right)'$ is the $(p+1) \times 1$ vector of parameter functions

$d\mathbf{M}(t)$ is the $n \times 1$ vector of martingales (error terms) each with mean zero.

It follows from (6) and from the theory of least square estimation that if $\mathbf{X}(t)$ is of full rank, that is $[\mathbf{X}(t)]'\mathbf{X}(t)$ is non singular, then the ordinary least squares estimator of $d\mathbf{B}(t)$ is

$$d\widehat{\mathbf{B}}(t) = \left[(\mathbf{X}(t))'\mathbf{X}(t)\right]^{-1}(\mathbf{X}(t))'d\mathbf{N}(t) \tag{7}$$

If $\mathbf{X}(t)$ is not of full rank, then $d\mathbf{B}(t)$ is not estimable unless some constraint is imposed. However, most of current statistical packages have built-in routines to deal with matrices that are not of full rank and provide robust estimates of model parameters. The estimator $\widehat{\mathbf{B}}(t)$ obtained by integrating both sides of Equation (7) with respect to t is

$$
\begin{aligned}
\widehat{\mathbf{B}}(t) &= \int_0^t \left[(\mathbf{X}(t))'\mathbf{X}(t)\right]^{-1}(\mathbf{X}(t))'d\mathbf{N}(t) \\
&= \sum_{t_j \leq t} \left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1}(\mathbf{X}(t_j))'\mathbf{y}_j
\end{aligned} \tag{8}
$$

where $\mathbf{y}_j$ is $n \times 1$ vector of zeros except the jth component equals to unit if the jth individual observes an event at time t_j (Aalen et al. 2008, p. 158; Hosmer and Royston 2002). Furthermore, the variance-covariance matrix of $\widehat{\mathbf{B}}(t)$ is

$$\mathrm{Var}\left[\widehat{\mathbf{B}}(t)\right] = \sum_{t_j \leq t} \left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1}(\mathbf{X}(t_j))'\mathbf{D}(t_j)\mathbf{X}(t_j)\left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1} \tag{9}$$

where $\mathbf{D}(t_j)$ is an $n \times n$ diagonal matrix with elements y_j on the main diagonal (Aalen et al. 2008, p. 158; Hosmer and Royston 2002). The derivation of results (9) from (8) is easy to understand. In fact if two random vectors of variables $\mathbf{X}$ and $\mathbf{Y}$ are linked by $\mathbf{Y} = \mathbf{A}\mathbf{X}$, where $\mathbf{A}$ is a matrix, then

$$\mathrm{Var}(\mathbf{Y}) = \mathbf{A}\,\mathrm{Var}(\mathbf{X})\,\mathbf{A}'$$

(Mulaik 2009).

As described in Hosmer and Royston (2002), if the vector of cumulative parameter coefficients at time t is estimated by (8), and its variance-covariance matrix by (9), then the estimator of the model vector of parameters at time t_j is

$$\hat{\boldsymbol{\beta}}(t_j) = \left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1}(\mathbf{X}(t_j))'\mathbf{y}_j \tag{10}$$

and

$$\mathrm{Var}\left[\widehat{\boldsymbol{\beta}}(t_j)\right] = \left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1}(\mathbf{X}(t_j))'\mathbf{D}(t_j)\mathbf{X}(t_j)\left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1}. \tag{11}$$

Aalen et al. (2008, p. 159) showed that the cumulative parameter function estimator $\widehat{\mathbf{B}}(t)$ has approximately a multivariate normal distribution around its true value $\mathbf{B}(t)$,

with the variance-covariance matrix expressed in (9). Therefore, the $100(1-\alpha)\%$ confidence interval for the kth cumulative parameter functions $B_k(t)$ is expressed by

$$\widehat{B}_k(t) \pm z_{\frac{\alpha}{2}} \sqrt{\hat{\sigma}_{kk}(t)} \tag{12}$$

where $\hat{\sigma}_{kk}(t)$ is the kth diagonal element of the variance-covariance matrix expressed in the Equation (9). To test that a covariate X_k has no significant effect on the hazard function given in model (3), Aalen et al. (2008) formulated the null and alternative hypotheses in the usual way as follows

$$H_0 : \beta_k(t) = 0, \ \forall \, t \in [0, t_0]$$

versus

$$H_1 : \beta_k(t) > 0 \text{ or } \beta_k(t) < 0$$

where t_0 is a suitably chosen time point, but often t_0 is the upper limit of the study time interval. If H_0 is true, then the increment $\Delta\hat{B}_k(t_j)$ at time t_j of the cumulative parameter function given in (8) tends to fluctuate around zero (Aalen et al. 2008). Under the alternative hypothesis $H_1 : \beta_k(t) > 0$, the increment $\Delta\hat{B}_k(t_j)$ tends to be positive while under $H_1 : \beta_k(t) < 0$, the increment $\Delta\hat{B}_k(t_j)$ tends to be negative. Furthermore if $\widehat{B}_k(t)$ approximately follows a straight line, then $\beta_k(t)$ is constant. The test described above is helpful when the estimated cumulative parameter functions are plotted against time. However, a quantitative measure of significance may be needed to assess the magnitude of significance. Hosmer and Royston (2002) advised to proceed as follows. Consider model (3) and assume that there is a need to test the null hypothesis

$$H_0 : \beta_k(t_j) = 0 \text{ for all } k \text{ with } k = 0, 1, \ldots, p. \tag{13}$$

Hosmer and Royston (2002) stated that the $(p+1)$ statistics for the above hypothesis are obtained from the components of the vector

$$\hat{\mathbf{u}} = \sum_{t_j} \mathbf{K}_j \hat{\boldsymbol{\beta}}(t_j) \tag{14}$$

where $\hat{\boldsymbol{\beta}}(t_j)$ given by (10) is the vector of estimators of the parameter coefficients for model (3), and $\mathbf{K}_j$ is a $(p+1) \times (p+1)$ diagonal matrix of weights. Four types of weights can be used.

- Weights 1: $\mathbf{K}_j = \text{diag}(1)$, that is $\mathbf{K}_j$ is a diagonal matrix with each element of the main diagonal equals to unit.
- Weights 2: $\mathbf{K}_j = \text{diag}(n_j)$ where n_j is the number of individuals at risk at time t_j.
- Weights 3: $\mathbf{K}_j = \text{diag}\left[\widehat{S}_{KM}(t_{j-1})\right]$ where $\widehat{S}_{KM}(t_{j-1})$ is the Kaplan–Meier estimate of the survival function at time t_{j-1} for $j = 2, 3, \ldots$ and $\mathbf{K}_1 = \text{diag}\left[\widehat{S}_{KM}(t_0) = 1\right]$.
- Weights 4: $\mathbf{K}_j = \text{diag}\left[\widehat{S}_{KM}(t_{j-1}) / \text{se}(\hat{\beta}_{kk}(t_j))\right]$ where $\hat{\beta}_{kk}(t_j)$ is the kth diagonal element of the variance-covariance matrix (11). Hence, $\mathbf{K}_j$ is a diagonal matrix whose main diagonal elements are the ratio of the Kaplan–Meier estimates of the survival function at time t_{j-1} and the standard error of the Aalen estimate of the parameter function of interest at time t_j.

To completely define the test statistic to use, the estimator of the variance-covariance matrix of $\hat{\mathbf{u}}$ given in (14) is obtained from the symmetric matrix $\mathbf{K}_j$ and the variance of $\hat{\beta}(t_j)$ given by (11). Hence,

$$
\begin{aligned}
\widehat{\text{Var}}(\hat{\mathbf{u}}) &= \sum_{t_j} \mathbf{K}_j \text{Var}\left[\hat{\boldsymbol{\beta}}(t_j)\right] \mathbf{K}_j' \\
&= \sum_{t_j} \mathbf{K}_j \left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1} (\mathbf{X}(t_j))'\mathbf{D}(t_j)\mathbf{X}(t_j)\left[(\mathbf{X}(t_j))'\mathbf{X}(t_j)\right]^{-1} \mathbf{K}_j'.
\end{aligned}
\tag{15}
$$

Hence, the test statistic for H_0 given in (13) is

$$
z_{u_k} = \frac{\hat{u}_k}{\text{se}(\hat{u}_k)}
\tag{16}
$$

where $\hat{u}_k$ is the kth element of $\hat{\mathbf{u}}$ given in (14) and $\text{se}(\hat{u}_k)$ is the square root of the kth diagonal element of $\widehat{\text{Var}}(\hat{\mathbf{u}})$ given in (15). Hosmer and Royston (2002) pointed out that the statistic z_{u_k} in (16) approximately follows the standard normal distribution.

2.5. Dataset

Among over 6000 active and inactive cryptocurrencies, we recorded 500 cryptocurrencies whose information on the variables of interest is available. The details of other cryptoccurrencies is not available, and for many of them, the white papers are not decentralized. The covariates of interest are described in Table 2.

Table 2. Description of variables of interest.

Variable	Description	Code/Value/Unit
Type	The type of a cryptocurrency	0 = token 1 = coin
Blockchain	Indicator of the type of blockchain of a cryptocurrency	0 = other 1 = Ethereum 2 = Standalone
Series	Indicator on when a cryptocurrency was released	0 = series 1 (2009–2013) 1 = series 2 (2013–2017) 2 = series 3 (2017–2021)
Mining	Indicator on whether a cryptocurrency is minable or not	0 = not minable 1 = minable
Region	Region in which a cryptocurrency' s'headquarters are based	0 = unknown 1 = South America 2 = Oceania 3 = North America 4 = Europe 5 = Asia 6 = Africa

The Kaplan–Meir estimation of the overall survival function (Figure 1) at the end of the study time is 75%, or equivalently, the rate of death at the end of the study time is 25%. The time is in years with origin fixed in 2009.
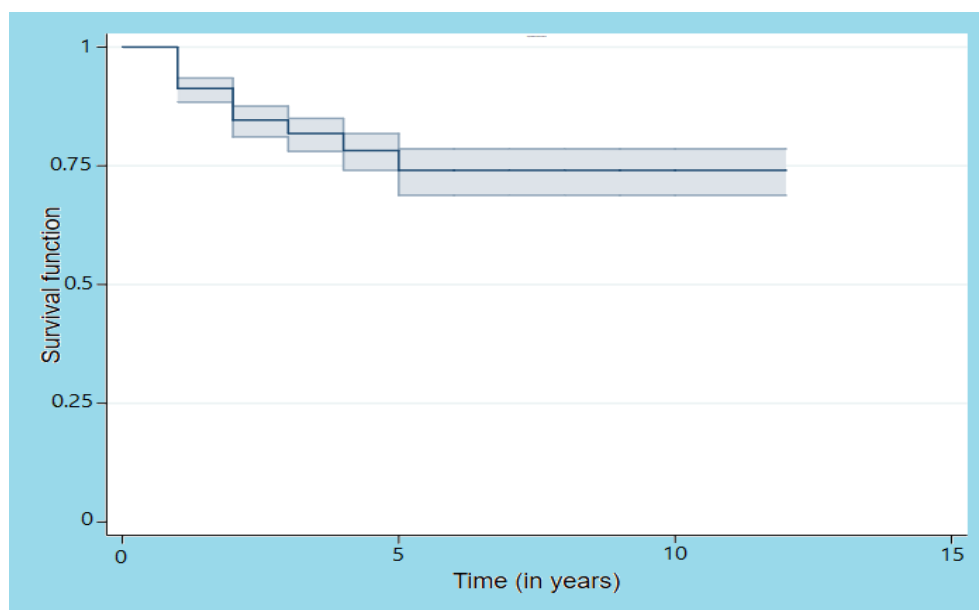
Figure 1. Kaplan–Meier estimates and 95% confidence limits of the survival function for the cryptocurrency data.

The survival outcomes of the levels of the variables can be compared graphically by using the Kaplan–Meier estimation of the survival function per group of covariate. The plots are displayed in Figure 2. The log-rank and Wilcoxon test statistics are summarized in Table 3.

Figure 2a suggests that the survival outcome is significantly better for tokens as compared to coins ($\chi^2 = 42.02$; $p < 0.001$). Figure 2b indicates a relatively better survival outcome for cryptocurrencies under Ethereum ($\chi^2 = 66.90$; $p < 0.001$). Figure 2d suggests a significant high risk as a cryptocurrency is minable ($\chi^2 = 32.22$; $p < 0.001$). Among the levels of the variable *Series*, cryptocurrencies released in the series 2009–2013 present significantly better survival outcomes ($\chi^2 = 8.04$; $p = 0.018$) as Figure 2d indicates, this is the same for South American cryptocurrencies ($\chi^2 = 132.03$; $p < 0.001$) as Figure 2e shows of the variable *Region*.

Table 3. Log-rank and Wilcoxon test statistics.

Variable	Log-Rank χ^2 Test Statistic (*p*-Value)	Wilcoxon χ^2 Test Statistic (*p*-Value)
Type	42.02 ($p < 0.001$)	35.45 ($p < 0.001$)
Blockchain	66.90 ($p < 0.001$)	55.62 ($p < 0.001$)
Mining	32.22 ($p < 0.001$)	22.34 ($p < 0.001$)
Series	13.10 (0.001)	8.04 (0.018)
Region	153.49 ($p < 0.001$)	132.03 ($p < 0.001$)

The log-rank test for comparison is suitable for comparing levels of covariates that obey the proportional hazard assumption (PHA); these are covariates *Type*, *Blockchain* and *Mining* for which the plots are approximately parallel. Wilcoxon test is suitable in comparing the levels of covariates *Series* and *Region* for which corresponding plots cross each other, leading to the violation of the PHA.
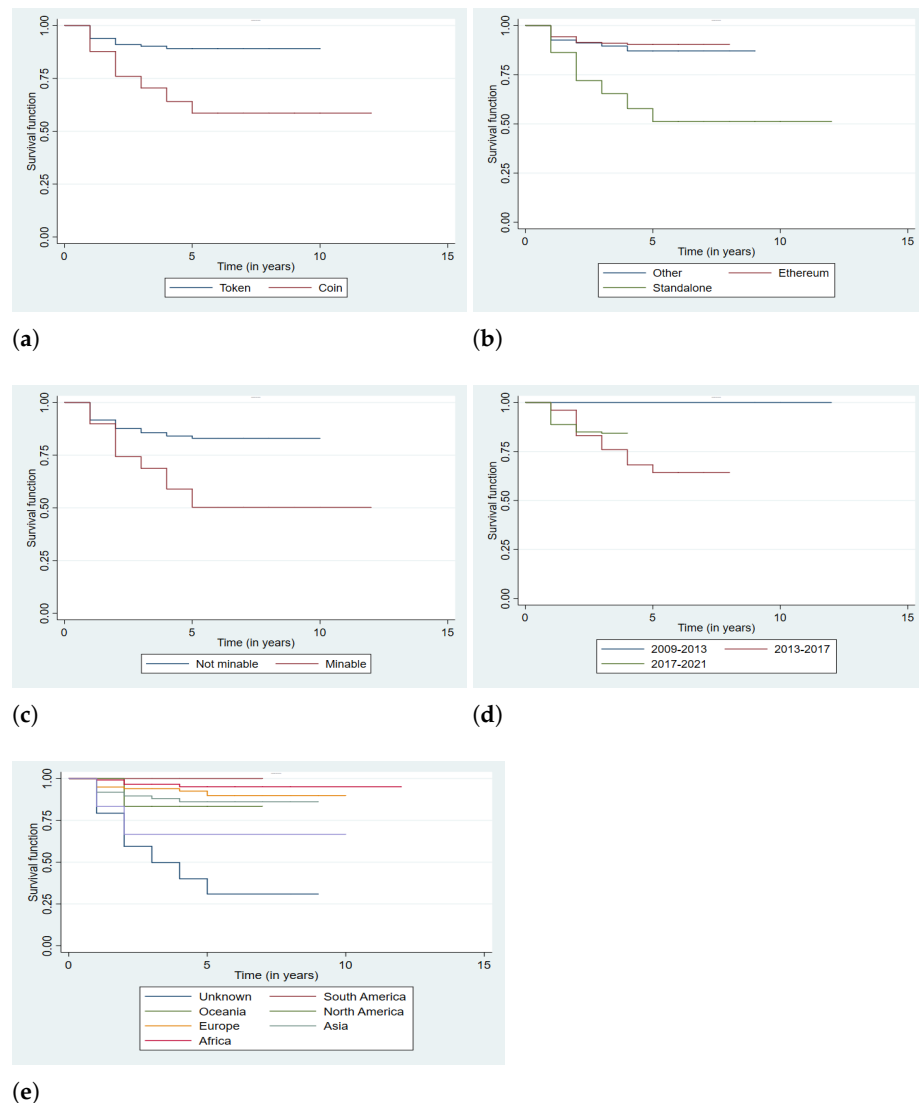
Figure 2. Plots of the Kaplan–Meier estimates of the survival function for levels of covariates from 2009 to 2021. (**a**) Type, (**b**) Blockchain, (**c**) Mining, (**d**) Series, (**e**) Region.

3. Results and Interpretation

3.1. Cox Proportional Hazards Model (CPHM)

Unlike Kaplan–Meier estimation, which treats one variable at a time, the the CPHM makes inference by considering several variables at a time.

Table 4 presents the estimated hazard ratios based on all the covariates.

The model suggests that the risk of death of cryptocurrencies under standalone blockchain is 2.771 (95% CI: 1.221; 6.288) times that of cryptocurrencies under the referral blockchain.

The significance for covariate *Region* is observed in North America, Europe and Asia. I all these regions the survival outcomes are better than the referral level. The CPHM suggests that the risk of death of cryptoccurrencies from unknown region is 6.757 (95% CI: [3.106; 14.706]) times that of cryptocurrencies based in North America. Such risk is 4.348 (95% CI: [2.439; 7.752]) times that of cryptocurrencies based in Europe and 13.333 (95% CI: [5.181; 34.483]) times that of cryptocurrencies based Asia.

Table 4. Cox Proportional hazards model for all covariates.

Covariate (Reference)	Level	Haz. Ratio	Std. Err.	z	P > z	95% Conf. Int.
Type (Token)	Coin	0.737	0.350	−0.640	0.521	[0.290; 1.872]
Blockchain (Other)	Ethereum	0.798	0.382	−0.470	0.638	[0.312; 2.041]
	Standalone	2.771	1.159	2.440	0.015	[1.221; 6.288]
Series (Series 1)	Series 2	3.270	0.230	2.840	0.671	[1.452; 3.215]
	Series 3	2.870	0.952	1.450	0.424	[1.413; 3.624]
Mining (Not minable)	Minable	0.855	0.208	−0.640	0.520	[0.530; 1.378]
Region (Unknown)	South America	0.091	0.396	−0.562	0.741	[0.067; 0.201].
	Oceania	0.402	0.414	−0.880	0.377	[0.053; 3.031]
	North America	0.148	0.059	−4.800	$p < 0.001$	[0.068; 0.322]
	Europe	0.230	0.068	−4.980	$p < 0.001$	[0.129; 0.410]
	Asia	0.075	0.036	−5.370	$p < 0.001$	[0.029; 0.193]
	Africa	0.755	0.554	−0.380	0.702	[0.179; 3.182]

*3.2. Aalen Additive Hazards Model (AAHM)*

Hosmer and Royston (2002) designed a STATA code based on an ado file for analysing survival data using the AAHM.

Figure 3a displays the cumulative parameter function with its 95% confidence limits for cryptocurrencies under Ethereum blockchain. The plot oscillates around the zero line, and the plot of 95% confidence limits are on either sides of the zero line. This indicates that the slope may be zero at some time value, and the risk of cryptocurrencies under Ethereum may not be significantly different from that of the referral group. The same AAHM results are observed for the crypto-coins (Figure 3c) and minable cryptocurrencies (Figure 3d).

Figure 3b indicates that the cumulative parameter function for cryptocurrencies under standalone blockchain is positive, and so is the major part of the 95% confidence limits. This suggests that the risk of such cryptocurrencies is higher than that of the referral group. The same observation occurred for cryptocurrencies released from the year 2013 as Figure 3e,f show.

Figure 4 gives the plots of the cumulative parameter functions and their 95% confidence limits for levels of the variable *Region*. The cumulative parameter functions with their 95% confidence limits are negative for levels *South America* (Figure 4a), *North America* (Figure 4c), *Europe* (Figure 4d) and *Asia* (Figure 4e). The pattern for level *Oceania* is negative together with the major parts of its confidence limits (Figure 4b). This suggests a relatively higher risk of cryptocurrencies of the referral region. The pattern for level *Africa* (Figure 4f) oscillates around the zero line, and the plot of 95% confidence limits are on either sides of the zero line. This indicates that the slope may be zero at some time value, and the risk of African cryptocurrencies may not be significantly different from that of the referral region.

Table 5 displays the results of the test statistics of the AAHM for all the covariates using the four types of weights. All the tests are against the difference of the levels of covariate *Mining*. The test based on type 4 weights suggests the moderate evidence of difference of the levels of covariate *Type* and the overwhelming difference of levels of covariates *Blockchain*, *Series* and *Region*. The overwhelming difference for levels of covariate *Series* and some levels of covariate *Region* is also noticed by test 1, 2 and 3.
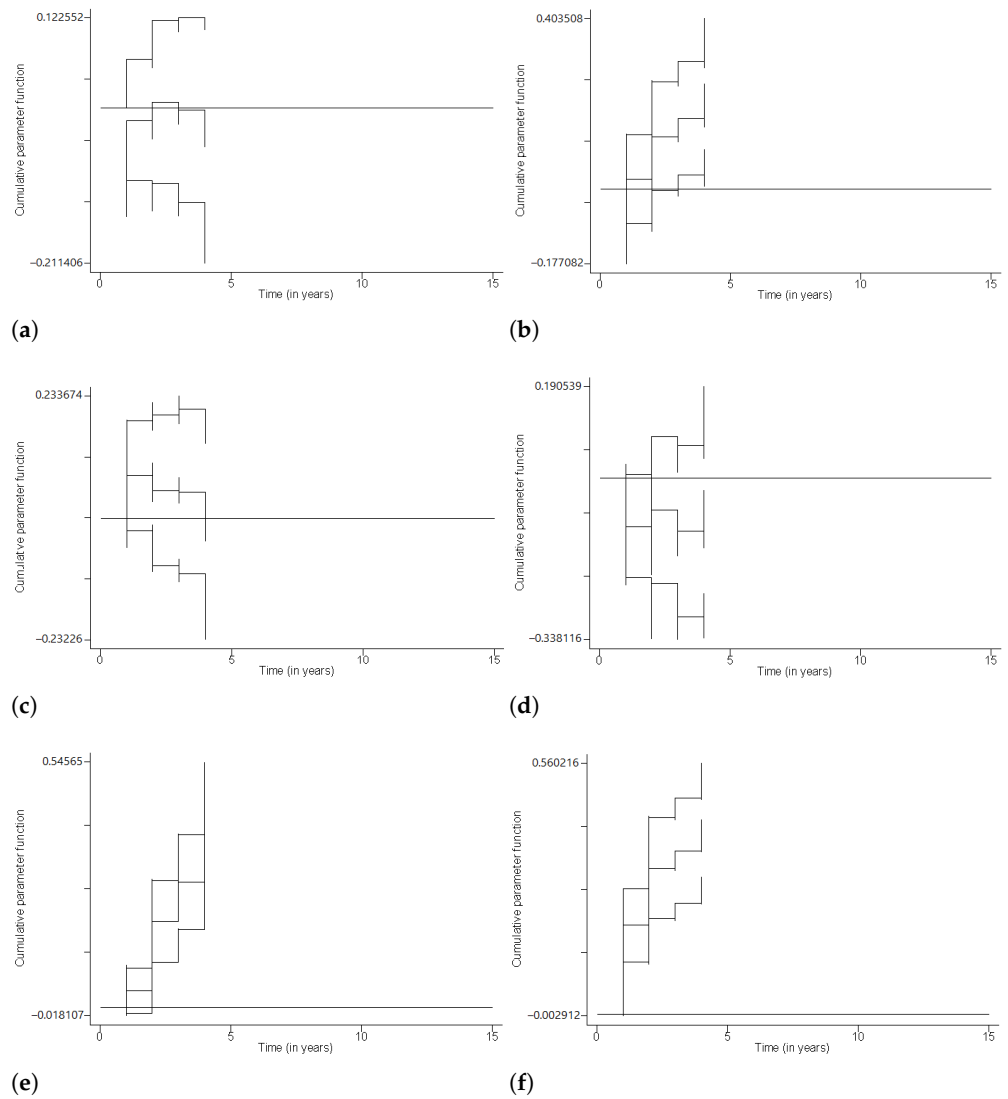
Figure 3. AAHM for levels of covariates *Type*, *Blockchain*, *Mining* and *Series*. (**a**) Ethereum, (**b**) Standalone, (**c**) Coin, (**d**) Minable, (**e**) Series 2 (2013–2017), (**f**) Series 3 (2017–2021).
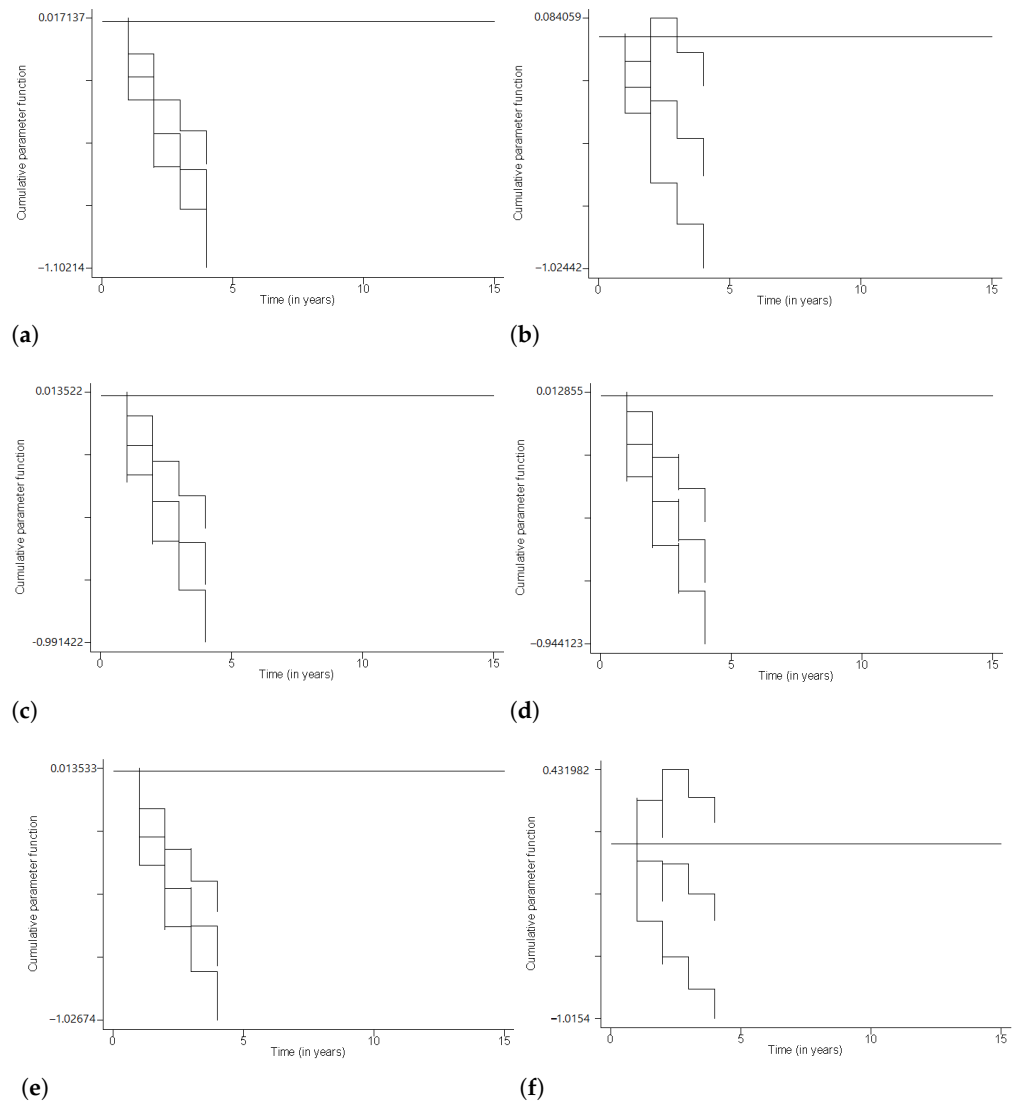
Figure 4. AAHM for levels of covariate *Region*. (**a**) South America, (**b**) Oceania, (**c**) North America, (**d**) Europe, (**e**) Asia, (**f**) Africa.

Table 5. Tests for significance of covariates.

Covariate (Reference)	Level	Test 1		Test 2		Test 3		Test 4	
		z	P	z	P	z	P	z	P
Type (Token)	Coin	−0.477	0.634	0.004	0.997	−0.385	0.700	−2.229	0.026
Blockchain (Other)	Ethereum	−0.451	0.652	−0.437	0.662	−0.510	0.610	7.230	*p* < 0.001
	Standalone	3.098	0.002	2.712	0.007	2.907	0.004	5.196	*p* < 0.001
Series (Series 1)	Series 2	6.114	*p* < 0.001	6.052	*p* < 0.001	6.090	*p* < 0.001	5.558	*p* < 0.001
	Series 3	6.519	*p* < 0.001	6.524	*p* < 0.001	6.526	*p* < 0.001	7.719	*p* < 0.001
Mining (Not minable)	Minable	−0.324	0.746	−0.754	0.451	−0.448	0.654	0.990	0.322
Region (Unknown)	South America	−7.367	*p* < 0.001	−7.784	*p* < 0.001	−7.487	*p* < 0.001	−7.365	*p* < 0.001
	Oceania	−2.956	0.003	−2.935	0.003	−2.984	0.003	−7.153	*p* < 0.001
	North America	−5.895	*p* < 0.001	−6.312	*p* < 0.001	−5.985	*p* < 0.001	−6.110	*p* < 0.001
	Europe	−5.332	*p* < 0.001	−5.661	*p* < 0.001	−5.403	*p* < 0.001	−5.497	*p* < 0.001
	Asia	−6.327	*p* < 0.001	−6.837	*p* < 0.001	−6.471	*p* < 0.001	−6.353	*p* < 0.001
	Africa	−1.468	0.142	−1.339	0.181	−1.418	0.156	−7.350	*p* < 0.001

4. Conclusions

This paper used survival regression models for analysing the risk to death of the cryptocurrencies from 2009 to Q2 2021. The dataset is a sample of 500 cryptocurrencies for which a correct information on the covariates of interest were found. The exploration was conducted using the Kaplan–Meier estimation of the survival function.

The Cox Proportional Hazards Model (CPHM) also was used and suggested a relatively higher risk for cryptocurrencies under standalone blockchain. This result was also found by the plot of the survival functions of the levels of blockchain covariate. It was found by the CPHM that cryptocurrencies released in the series 2013–2017 are at a high risk as compared to those of 2009–2013 release. The CPHM also suggested that cryptocurrencies for which headquarters are unknown are at a relatively higher risk. The results of the Aalen Additive Hazards Models (AAHM) showed that unlike other covariates, the levels of covariate *Mining* are not significantly different.

Among more than 6000 active and inactive cryptocurrencies, this paper considered 500 cryptocurrencies for which the information on the variables of interest were easily found. The study still needs improvement by considering a relatively bigger sample size. Apart from considering a big sample in future research, re-sampling may also improve the measurement of the standard errors and then evaluate the accuracy of the results found in this paper.

References

Aalen, Odd Olai, Ørnulf Borgan, and Håkon Kristian Gjessing. 2008. *Survival and Event History Analysis*. New York: Springer.

Angela, Scott-Briggs. 2016. Ten types of digital currencies and how they work. *Online Trading: Free Introductory eBook* 2016: 24.

Azimli, Asil. 2020. The impact of COVID-19 on the degree of dependence and structure of risk-return relationship: A quantile regression approach. *Finance Research Letters* 36: 101648. [CrossRef] [PubMed]

Blundell-Wignall, Adrian. 2014. The Bitcoin question: Currency versus trust-less transfer technology. In *OECD Working Papers on Finance, Insurance and Private Pensions, No 37*. Paris: OECD Publishing.

Breslow, Norman Edward. 1974. Covariance analysis of censored survival data. *Biometrics* 30: 89–99. [CrossRef] [PubMed]

Chan, Stephen, Jeffrey Chu, Saralees Nadarajah, and Joerg Osterrieder. 2017. A statistical analysis of cryptocurrencies. *Journal of Risk and Financial Management* 10: 12. [CrossRef]

Collet, David. 2003. *Modeling Survival Data in Medical Research*, 2nd ed. London: Chapman & Hall.

Cox, David Roxbee. 1972. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* 34: 187–220.

De Best, Raynor. 2022. Number of cryptocurrencies worldwide. *Statista* 1: 2022.

Efron, Bradley. 1977. The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association* 72: 557–65. [CrossRef]

Gatabazi, Paul, and Gaëtan Kabera. 2015. *Survival Analysis and Its Stochastic Process Approach with Application to Diabetes Data*. Johannesburg: University of Johannesburg.

Gatabazi, Paul, Jules Clement Mba and Edson Pindza. 2019a. *Analysis of Cryptocurrencies Adoption Using Fractional Grey Lotka-Volterra Models*. Johannesburg: University of Johannesburg.

Gatabazi, Paul, Jules Clement Mba, and Edson Pindza. 2019b. Fractional Grey Lotka-Volterra model with application to cryptocurrencies adoption. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29: 073116. [CrossRef] [PubMed]

Gatabazi, Paul, Jules Clement Mba, and Edson Pindza. 2019c. Modeling cryptocurrencies transaction counts using variable-order fractional grey Lotka-Volterra dynamical system. *Chaos Solitons & Fractals* 127: 283–90.

Gatabazi, Paul, Jules Clement Mba, Edson Pindza, and Coenraad Labuschagne. 2019d. Grey Lotka-Volterra model with application to cryptocurrencies adoption. *Chaos, Solitons & Fractals* 122: 47–57.

Gatabazi, Paul, Melesse Sileshi Fanta, and Shaun Ramroop. 2019e. Resampled Cox proportional hazards model for the infant mortality at the Kigali University Teaching Hospital. *The Open Public Health Journal* 12: 136–44. [CrossRef]

Gatabazi, Paul, Melesse Sileshi Fanta, and Shaun Ramroop. 2018. Multiple events model for the infant mortality at Kigali University Teaching Hospital. *The Open Public Health Journal* 11: 464–73. [CrossRef]

Gatabazi, Paul, Melesse Sileshi Fanta, and Shaun Ramroop. 2020a. Comparison of three classes of marginal risk set model in predicting infant mortality among newborn babies at Kigali University Teaching Hospital, Rwanda, 2016. *BMC Pediatrics* 20: 62. [CrossRef] [PubMed]

Gatabazi, Paul, Melesse Sileshi Fanta, and Shaun Ramroop. 2020b. *Some Statistical Methods in Analysis of Single and Multiple Events with Application to Infant Mortality Data*. KwaZulu Natal: University of KwaZulu Natal.

Hendrickson, Joshua, Thomas Hogan, and William Luther. 2016. The political economy of Bitcoin. *Economics Inquiry* 54: 925–39. [CrossRef]

Hosmer, David, Stanley Lemeshow, and Susanne May. 2008. *Regression Modeling of Time-to-Event Data*, 2nd ed. Haboken: John Wiley & Sons Inc.

Hosmer, David and Patrick Royston. 2002. Using Aalen's linear hazards model to investigate time-varying effects in the proportional hazards regression model. *The STATA Journal* 2: 331–50. [CrossRef]

Klein, John, and Melvin Moeschberger. 2003. *Techniques for Censored and Truncated Data*, 2nd ed. New York: Springer.

Mba, Jules Clement, and Qing-Guo Wang. 2019. *Multi-Period Portfolio Optimization: A Differential Evolution Copula-Based Approach*. Johannesburg: University of Johannesburg.

Mulaik, Stanley. 2009. *Foundations of Factor Analysis*, 2nd ed. Boca Raton: Chapman & Hall/CRC.

Thul, Joaquin. 2021. El Salvador legalises Bitcoin: Hype or breakthrough? *Infocus Macro Comment* 9: 2021.

Urquhart, Andrew. 2016. The inefficiency of Bitcoin. *Economics Letters* 148: 80–82. [CrossRef]

Wu, Ke, Spencer Wheatley, and Didier Sornette. 2018. Classification of cryptocurrency coins and tokens by the dynamics of their market capitalizations. *Royal Sosiety Open Science* 5: 180381. [CrossRef]

Xiao, Hui, Xiong Xiong, and Weiwei Chen. 2021. Introduction to the special issue on impact of COVID-19 and cryptocurrencies on the global fnancial market. *Financial Innovation* 7: 27. [CrossRef]

Youssef, Manel, Khaled Mokni, and Ahdi Noomen Ajmi. 2021. Dynamic connectedness between stock markets in the presence of the COVID-19 pandemic: Does economic policy uncertainty matter? *Financial Innovation* 7: 13. [CrossRef] [PubMed]