Crypto Coins and Credit Risk: Modelling and Forecasting their Probability of Death

Fantazzini, Dean

2022

Online at https://mpra.ub.uni-muenchen.de/113744/
MPRA Paper No. 113744, posted 13 Jul 2022 00:26 UTC
Abstract

This paper examined a set of over two thousand crypto-coins observed between 2015 and 2020 to estimate their credit risk by computing their probability of death. We employed different definitions of dead coins, ranging from academic literature to professional practice, alternative forecasting models, ranging from credit scoring models to machine learning and time series-based models, and different forecasting horizons. We found that the choice of the coin death definition affected the set of the best forecasting models to compute the probability of death. However, this choice was not critical, and the best models turned out to be the same in most cases. In general, we found that the cauchit and the zero-price-probability (ZPP) based on the random walk or the Markov Switching-GARCH(1,1) were the best models for newly established coins, whereas credit scoring models and machine learning methods using lagged trading volumes and online searches were better choices for older coins. These results also held after a set of robustness checks that considered different time samples and the coins’ market capitalization.

Keywords: Bitcoin, Crypto-assets, Crypto-currencies, Credit risk, Default Probability, Probability of Death, ZPP, Cauchit, Logit, Probit, Random Forests, Google Trends.

JEL classification: C32, C35, C51, C53, C58, G12, G17, G32, G33.
Introduction

Cryptoasset research has become a hot topic in the field of finance: for example (and to name just a few) [Antonopoulos, 2014] describes the technical foundations of bitcoin and other cryptographic currencies, from cryptography basics, such as keys and addresses, to the data structures, network protocols and the consensus mechanism, while [Narayanan et al., 2016] provide a comprehensive introduction to digital currencies. [Burniske and Tatar, 2018] discuss a general framework for investigating and valuing cryptoassets, [Brummer, 2019] focuses on the legal, regulatory, and monetary issues of the whole crypto ecosystem, [Fantazzini, 2019] discusses the instruments needed to analyze cryptocurrencies markets and prices, while [Schar and Berentsen, 2020] provide a general introduction to cryptocurrencies and blockchain technology for practitioners and students.

The increasing number of traded crypto assets\(^1\) and the repeated cases of hacks, scams, and projects’ failures have made the topic of crypto assets risk a compelling issue, see [Fantazzini and Zimin, 2020] and references therein. A cryptocurrency does not have debt and it cannot default in a classical sense\(^2\), but its price can crash quickly due to a hack, a scam, or other problems that can make its further development no more viable. [Fantazzini and Zimin, 2020] showed this kind of risk is not a market one and proposed a new definition of credit risk for crypto-coins based on their “death”, that is a situation when their price drop significantly and a coin becomes illiquid.

We remark that there is not a unique definition for a dead coin, neither in the professional literature\(^3\)\(^4\) nor in the academic literature, see [Feder et al., 2018], [Grobys and Sapkota, 2020] and [Schmitz and Hoffmann, 2020]. Moreover, even when a coin is considered dead, it may still show some minimal trading volumes, either due to the possibility to recover a small amount of the initial investment, or simply to bet on its possible revamp. In this regard, a coin can be easily revamped by writing new code or simply by updating the previous old code, thus involving much less time and resources than traditional bankrupt firms, see [Sid, 2018] for an example. Therefore, the “death” state for a coin can be only a temporary state rather than a permanent one.

Despite the presence of thousands of dead coins and a yearly increase in 2021 of more than 30%\(^1\)

---

\(^1\)At the end of December 2021, almost 15000 crypto assets were listed on Coinmarketcap.com. CoinMarketCap is the main aggregator of cryptocurrency market data, and it has been owned by the crypto exchange Binance since April 2020, see https://crypto.marketswiki.com/index.php?title=CoinMarketCap for more details.

\(^2\)Lansky, (2018, p.19) formally defined a cryptocurrency as a system that satisfies these six conditions: “1) The system does not require a central authority, its state is maintained through distributed consensus. 2) The system keeps an overview of cryptocurrency units and their ownership. 3) The system defines whether new cryptocurrency units can be created. If new cryptocurrency units can be created, the system defines the circumstances of their origin and how to determine the ownership of these new units. 4) Ownership of cryptocurrency units can be proved exclusively cryptographically. 5) The system allows transactions to be performed in which ownership of the cryptographic units is changed. A transaction statement can only be issued by an entity proving the current ownership of these units. 6) If two different instructions for changing the ownership of the same cryptographic units are simultaneously entered, the system performs at most one of them.”

\(^3\)https://www.investopedia.com/news/crypto-carnage-over-800-cryptocurrencies-are-dead/

\(^4\)https://www.coinopsy.com/dead-coins/.
This topic has been barely examined in the academic literature. [Feder et al., 2018] were the first to propose a formal definition of dead coin, while [Schmitz and Hoffmann, 2020] suggested some simplified procedures to identify a dead coin for portfolio management. [Fantazzini and Zimin, 2020] and [Grobys and Sapkota, 2020] were the first (and so far only) to propose models to predict cryptocurrency defaults/deaths.

This paper aims to forecast the probability of death of a crypto coin using different definitions of dead coins, ranging from academic literature to professional practice, and different forecasting horizons. To reach the paper’s objective, we first employed a set of models to forecast the probability of death, including credit scoring models, machine learning models, and time-series methods based on the Zero Price Probability (ZPP) model by [Fantazzini et al., 2008], which is a methodology to compute the probabilities of default using only market prices. Recent papers by [Su and Huang, 2010], [Li et al., 2016], [Dalla Valle et al., 2016], and [Fantazzini and Zimin, 2020] showed that ZPP models often outperform the competing models in terms of default probability estimation.

The second contribution of this paper is a forecasting exercise using a unique set of 2003 crypto coins that were active from the beginning of 2014 till the end of May of 2020. Our results show that the choice of the coin death definition can significantly affect the set of the best forecasting models to compute the probability of death. However, this choice is not critical, and the best models turned out to be the same in most cases. In general, we found that the cauchit and the zero-price-probability (ZPP) based on the random walk or the Markov Switching-GARCH(1,1) were the best models for newly established coins, whereas credit scoring models and machine learning methods using trading volumes and online searches are better choices for older coins.

The third contribution of the paper is a set of robustness checks to verify that our results also hold when considering different time samples and the coins’ market capitalization.

The paper is organized as follows: Section 2 briefly reviews the literature devoted to the credit risk of crypto-coins, while the methods proposed to model and forecast their probability of death are discussed in Section 3. The empirical results are reported in Section 4, while robustness checks are discussed in Section 5. Section 6 briefly concludes.

2 Literature review

The financial literature dealing with the credit risk involved in crypto coins is very limited and, at the time of writing this paper, only four papers examined the topic of dead coins, while only two of them proposed methods to forecast the probability of a coin death. We remark that when investing in a crypto

We will use the terms ‘probability of death’ and ‘probability of default’ interchangeably.
coin, there are two types of credit risks: the possibility that the coin “dies” and the price goes to zero (or close to zero), and the possibility that the exchange closes, taking most of its investors’ money with it. We focus here on the first type of risk, while the latter was examined in [Fantazzini and Calabrese, 2021] who considered a unique dataset of 144 exchanges, active from the first quarter of 2018 to the first quarter of 2021, to analyze the determinants surrounding the decision to close an exchange using credit scoring and machine learning techniques.

Currently, there is not a unique definition of dead coins, neither in the professional literature, nor in the academic literature: in the professional literature, some define dead coins as those whose value drops below 1 cent \(^6\), yet others stress, on top of that, no trading volume, no nodes running, no active community, and de-listing from (almost) all exchanges\(^7\).

[Feder et al., 2018] were the first to propose a formal definition of dead coin in the academic literature: they first define a ‘candidate peak’ as a day in which the 7-day rolling price average is greater than any value 30 days before or after. Moreover, to choose only those peaks with sudden jumps, they define a candidate as a peak only if it is greater than or equal 50% of the minimum value in the 30 days prior to the candidate peak, and if its value is at least 5% as large as the cryptocurrency’s maximum peak. Given these peak data, [Feder et al., 2018] consider a coin abandoned (= dead), if the daily average volume for a given month is less than or equal to 1% of the peak volume. Besides, if the currency is currently considered dead/abandoned but the average daily trading volume for a month following a peak is greater than 10% of the peak value, then [Feder et al., 2018] change the coin status to resurrected.

[Schmitz and Hoffmann, 2020] proposed a simplified version of the previous method by [Feder et al., 2018], and they suggested that a cryptocurrency can be classified as dead if its average daily trading volume for a given month is lower or equal to 1% of its past historical peak. Instead, a dead cryptocurrency is classified as ‘resurrected’ if this average daily trading volume reaches a value of more or equal to 10% of its past historical peak again\(^8\).

[Grobys and Sapkota, 2020] and [Fantazzini and Zimin, 2020] were the first (and so far only) to propose models to predict cryptocurrency defaults/deaths. [Grobys and Sapkota, 2020] examined a dataset of 146 Proof-of-Work-based cryptocurrencies that started trading before 2015 and followed their performance until December 2018, finding that about 60% of those cryptocurrencies died. They employed a model based on linear discriminant analysis to predict these defaults and found that it could predict most of the cryptocurrency bankruptcies, but it struggled to predict functioning cryptocurrencies. Predicting well the first category and poorly the second one is a well-known problem when using binary classification.

\(^6\)https://www.investopedia.com/news/crypto-carnage-over-800-cryptocurrencies-are-dead/  
\(^7\)https://www.coinopsy.com/dead-coins/  
\(^8\)Note that [Schmitz and Hoffmann, 2020] presented this method as the [Feder et al., 2018] approach when in reality the latter involves many more restrictions. The methodology used by Schmitz and Hoffmann (2020) in their empirical analysis is even more simplified, and it assumes that a coin is (temporarily) inactive if data gaps are present in its time series.
models. For this reason, model selection is usually based on loss functions such as the [Brier, 1950] score or the area under the receiver operating characteristic curve (AUC or AUROC) proposed by [Metz, 1978], [Metz and Kronman, 1980], and [Hanley and McNeil, 1982], instead of using the forecasting accuracy for each binary class. Another problematic issue with the analysis performed in [Grobys and Sapkota, 2020] is the need to use several coin-specific variable candidates that might serve as predictor variables: unfortunately, this kind of information is not available for most dead coins, and [Grobys and Sapkota, 2020] had to discard several variables to have a meaningful dataset. Moreover, considering the large number of scams and frauds regularly taking place among crypto assets, it is not advisable to take publicly available coin information at face value because it may be false. Besides, [Grobys and Sapkota, 2020] only performed an in-sample forecasting analysis, and they did not predict crypto-currencies that were not used to estimate their model. Unfortunately, there may be major differences between in-sample and out-of-sample forecasting performances, see [Hastie et al., 2009], [Giudici and Figini, 2009], and [Hyndman and Athanasopoulos, 2018] for a discussion at the textbook level.

[Fantazzini and Zimin, 2020] proposed a set of models to estimate the probability of death for a group of 42 crypto-currencies using the Zero Price Probability (ZPP) model by [Fantazzini et al., 2008], which is a methodology to compute the probabilities of default using only market prices, as well as credit scoring models and machine learning methods. Their empirical analysis showed that classical credit scoring models performed better in the training sample, whereas the models’ performances were much closer in the validation sample, with the simple ZPP computed using a random walk with drift performing remarkably well. The main limitation of the analysis performed by [Fantazzini and Zimin, 2020] is the very low number of coins used for backtesting (only 42), which can strongly limit the significance of their empirical evidence.

The past literature and professional practice highlighted that the dead coins collected in well-known online repositories such as coinopsy.com or deadcoins.com are indeed dead, but this fact represents (paradoxically) a problem. Unfortunately, the information set for the vast majority of these coins does not exist anymore because their technical information and historical market data are no more available. In simple terms, when a coin name is inserted in these repositories, it is too late to gain any valuable information for credit risk modelling and forecasting. It is for this reason that [Grobys and Sapkota, 2020] and [Fantazzini and Zimin, 2020] were forced to use small coin datasets in their analyses and to employ a limited set of variables to forecast these dead coins. Therefore, it makes more sense to employ the methods proposed by [Feder et al., 2018] and [Schmitz and Hoffmann, 2020] to detect dead coins, or the simple professional rule that defines a coin dead if its value drops below 1 cent. Even though there

---

9See section 5 in [Giudici and Figini, 2009] for a review.
10In-sample analysis is also known as training, while the out-of-sample analysis can be named as validation.
is still some marginal trading for the coins defined dead according to these rules, this is not a problem but an advantage because we can analyze them before they go into permanent (digital) oblivion.

Another issue that emerged from the literature review is the need to use indicators and methods that are robust to potential frauds and scams. As highlighted by [Fantazzini and Zimin, 2020], the lack of financial oversight for several crypto-based companies and exchanges means that coins prices can be subject to manipulations, pump-and-dump schemes and market frauds of various types, see [Gandal et al., 2018], [Wei, 2018], [Griffin and Shams, 2020], [Hamrick et al., 2021], and [Gandal et al., 2021] for more details about these unlawful acts.

3 Materials and Methods

We consider three approaches to forecast the probability of death of a large set of crypto coins: credit scoring models, machine learning, and time series methods. A review of the (large) literature on credit scoring models can be found in [Baesens and Van Gestel, 2009] and [Joseph, 2013], while for machine learning methods in finance we refer to [James et al., 2013], [De Prado, 2018] and [Dixon et al., 2020]. Time-series methods based on market prices to compute the probability of default of quoted stocks and Small and Medium Enterprises (SMEs) are discussed in [Fantazzini et al., 2008], [Su and Huang, 2010], [Li et al., 2016], [Dalla Valle et al., 2016], and [Jing et al., 2021], while their use with crypto coins is explored in [Fantazzini, 2019] and [Fantazzini and Zimin, 2020].

We first briefly review the main aspects of credit risk for cryptocurrencies. Secondly, we discuss a set of credit scoring and machine learning models that will be used in the empirical analysis. Then, time-series methods based on the ZPP originally proposed by [Fantazzini et al., 2008] as well as new variants are presented. Fourthly, we review several metrics to evaluate the estimated death probabilities. Finally, we also present the data used in our empirical analysis.

3.1 Credit risk for crypto-coins

In traditional finance, credit risk is defined as the gains and losses on a position or portfolio associated with the fulfillment (or not) of contractual obligations, while market risk as the gains and losses on the value of a position or portfolio that can take place due to the movements in market prices (such as exchange rates, commodity prices, interest rates, etc.), see [Basel Committee on Banking Supervision, 2009], [Hartmann, 2010] and references therein for more details. However, the [Basel Committee on Banking Supervision, 2009] highlighted that “the securitization trend in the last decade has diminished the scope for differences in measuring market and credit risk; as securitization transforms the latter into the former” (Basel Committee on Banking Supervision (2009), p.14). Besides, a large literature
showed that market and credit risk are driven by the same economic factors, see the special issue on the interaction of market and credit risk in the Journal of Banking and Finance in 2010 for more details.

[Fantazzini and Zimin, 2020] highlighted that the separation between market and credit risk becomes even more blurred when dealing with cryptocurrencies than in traditional finance. In simple terms, the credit risk for a crypto-coin is its “death”, a situation when its price falls significantly and a coin becomes illiquid. More formally, [Fantazzini and Zimin, 2020] define the “credit risk for cryptocurrencies as the gains and losses on the value of a position of a cryptocurrency that is abandoned and considered dead according to professional and/or academic criteria, but which can be potentially revived and revamped”.

Therefore, it follows that the differences between credit and market risk for cryptocurrencies are of quantitative and temporal nature, not qualitative because, if the financial losses and the technical problems are small, then we have a market event whereas, if the financial losses are too big and the technical problems cannot be solved, then we have a credit event and the cryptocurrency “dies” ([Fantazzini and Zimin, 2020]). Besides, the longer is the time horizon, the more probable are large losses and/or technical problems, so credit risk becomes more important. Once a credit event takes place, the development of the crypto-coin stops, and its price falls close to zero, or even to zero (if the lack of trading for several days or weeks is considered evidence of a zero price). However, trading may continue afterward for the reasons discussed in the introduction, that is for the possibility to recover a small amount of the initial investment, or simply to bet on its possible revamp.

More specifically, we employed three competing criteria to classify a coin as dead or alive in our work:

- the approach by [Feder et al., 2018]: first, a ‘candidate peak’ is defined as a day in which the 7-day rolling price average is greater than any value 30 days before or after. Moreover, to choose only those peaks with sudden jumps, a candidate is defined as a peak only if it is greater than or equal 50% of the minimum value in the 30 days prior to the candidate peak, and if its value is at least 5% as large as the cryptocurrency’s maximum peak. Given these peak data, [Feder et al., 2018] consider a coin abandoned (= dead), if the daily average volume for a given month is less than or equal to 1% of the peak volume. Besides, if the average daily trading volume for a month following a peak is greater than 10% of the peak value and that currency is currently abandoned, then [Feder et al., 2018] change the coin status to resurrected.

- the simplified [Feder et al., 2018] approach proposed by [Schmitz and Hoffmann, 2020]: a cryptocurrency can be classified as dead if its average daily trading volume for a given month is lower or equal to 1% of its past historical peak. Instead, a dead cryptocurrency is classified as ‘resurrected’

---

Note that this result is already known in the traditional financial literature because “the ratio of default and (normally distributed) market risk losses is proportional to the square-root of the holding period. Since the ratio goes to 0 as the holding period goes to 0, over short horizons market risk is relatively more important, while over longer horizons losses due to default become more important” ([Basel Committee on Banking Supervision, 2009], p. 16-17).
if this average daily trading volume reaches a value of more or equal to 10% of its past historical peak again.

- the professional rule that defines a coin dead if its value drops below 1 cent, and alive if its value rises above 1 cent.

### 3.2 Credit Scoring models and Machine Learning

Scoring models merge different variables into a quantitative score, which can be either interpreted as a probability of default (PD), or used as a classification system, depending on the model used. In the former case and considering our framework, a scoring model has the following form:

\[
P_{D_{i,t+T}} = \mathbb{P}(D_{i,t+T} = 1|D_{i,t} = 0; \mathbf{X}_{i,t}) = F(\beta' \mathbf{X}_{i,t})
\]

where \(P_{D_{i,t+T}}\) is the probability of death for the coin \(i\) over a period of time \(t + T\), given that it is alive at the time \(t\), and \(\mathbf{X}_{i,t}\) is a vector of regressors. If we use the logit model, or the probit model, or the cauchit model, \(F(\beta' \mathbf{X}_{i,t})\) is given by the logistic, standard normal, standard Cauchy -respectively-cumulative distribution function,

\[
F_{\text{Logit}}(\beta' \mathbf{X}_{i,t}) = \frac{1}{1 + e^{-(\beta' \mathbf{X}_{i,t})}}
\]

\[
F_{\text{Probit}}(\beta' \mathbf{X}_{i,t}) = \Phi(\beta' \mathbf{X}_{i,t}) = \int_{-\infty}^{(\beta' \mathbf{X}_{i,t})} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2} dz \tag{1}
\]

\[
F_{\text{Cauchit}}(\beta' \mathbf{X}_{i,t}) = \frac{1}{\pi} \left[ \tan^{-1}(\beta' \mathbf{X}_{i,t}) + \frac{\pi}{2} \right]
\]

The maximum likelihood method is usually used to estimate the parameters vector \(\beta\) in the equations (1), see [McCullagh and Nelder, 1989] for more details.

The logit and probit models are the widely used benchmarks for credit risk management, see [Fuertes and Kalotychou, 2006], [Rodriguez and Rodriguez, 2006], [Fantazzini and Figini, 2008], [Fantazzini and Figini, 2009], and references therein. The Cauchy distribution has heavier tails than the normal and logistic distributions, thus allowing more extreme values. As discussed in details by [Koenker and Yoon, 2009], the cauchit model can be used to model binary responses when observations occur for which the linear predictor is large in absolute value, indicating that the outcome is rather certain but yet the outcome is different. The cauchit model is more forgiving of these “outliers” than the logit or probit models. Besides, [Gündüz and Fokoué, 2017] shed some light on the theoretical reasons that explain the similar performance of four binary models (logit, probit, cauchit, and complementary log-log) in univariate settings. However, their simulation studies highlighted that the performance of the
four models in high dimensional spaces tends to depend on the internal structure of the input, with the cauchit being the model of choice under a high level of sparseness of the input space.

Machine Learning (ML) deals with the development of systems able to recognize complex patterns and make correct choices using a dataset already analyzed. Among the many methods available, we will use the Random Forest algorithm proposed by Ho (1995) and Breiman (2001), given its excellent past performances in forecasting binary variables, see [Hastie et al., 2009], [Barboza et al., 2017], [Moscatelli et al., 2020], and [Fantazzini and Calabrese, 2021] for more details. A Random Forest is an ensemble method consisting of a large number of decision trees, where a decision tree is similar to a reversed tree diagram with branches and leaves, where a choice is made at each step based on the value of a single variable, or a combination of several variables. In case of a classification problem, each leaf places an object either in one class or the other. A single decision tree can provide a poor classification and suffer from over-fitting and model instability. Random forests solve these problems by aggregating several decision trees into a so-called “forest”, where each tree is obtained by introducing a random component in their construction. More specifically, each decision tree in a forest is built using a bootstrap sample from the original data, where 2/3 of these data are used to build a tree, while the remaining 1/3 is used as a control set which is known as out-of-bag (OOB) data. Also, $m$ variables out of the original $n$ variables are randomly selected at each node of the tree, and the best split based on these $m$ variables is used to split the node. The random selection of variables at each node decreases the correlation among the trees in the forest so that the algorithm can deal with redundant variables and avoid model overfitting. Moreover, each tree is grown up to its maximum size and not pruned to maximize its instability, which is neutralized by the high number of trees created to have the “forest”. We remark that for a given $i$-th crypto-coin in the OOB control set, the forecasts are computed using a majority vote, which means that the probability of death is given by the proportion of trees voting for the death of coin $i$. This procedure is repeated for all observations in the control set, which leads to the computation of the overall OOB classification error.

3.3 Time-series methods

The Zero Price Probability (ZPP) was originally introduced in [Fantazzini et al., 2008] to compute the probabilities of default of traded stocks using only market prices $P_t$. This approach computes the market-implied probability $\mathcal{P}(P_\tau \leq 0)$ with $t < \tau \leq t + T$ using the fact that, for a traded stock (or a traded coin), the price $P_\tau$ is a truncated variable that cannot become less than zero. Therefore, the Zero-Price Probability is simply the probability that $P_\tau$ goes below the truncation level of zero. [Fantazzini et al., 2008] discussed in detail why the null price can be used as a default barrier.
The general estimation procedure of the ZPP for univariate time series is reported below:\footnote{\cite{Fantazzini and Zimin, 2020} proposed a multivariate approach to compute the ZPP of 42 coins. Given the very large dataset at our disposal, such an approach is not feasible in our case due to the curse of dimensionality. An extension of this methodology to large portfolios is left as an avenue for further research.}

1. Consider a generic conditional model for the differences of prices levels $X_t = P_t - P_{t-1}$ without the log-transformation:

$$X_t = \mu_t + \sigma_t z_t, \quad z_t \sim i.i.d f(0,1) \tag{2}$$

where $\mu_t$ is the conditional mean, $\sigma_t$ is the conditional standard deviation, while $z_t$ represents the standardized error.

2. Simulate a high number $N$ of price trajectories up to time $t + T$, using the estimated time series model (2) at step 1. We will compute the 1-day ahead, 30-day ahead, and 365-day ahead probability of death for each coin, that is $T = \{1, 30, 365\}$, respectively.

3. The probability of default/death for a crypto-coin $i$ is simply the ratio $n/N$, where $n$ is the number of times out of $N$ when the simulated price $P^k_t$ touched or crossed the zero barrier along the simulated trajectory:

$$PD_{i,t+T} = \frac{1}{N} \sum_{k=1}^{N} 1 \{ P^k_{\tau,i} \leq 0, \text{ for some } t < \tau \leq t + T \}$$

The previously cited literature dealing with the ZPP showed that the modelling of the conditional standard deviation $\sigma_t$ and the conditional distribution $f(\cdot)$ are the key elements affecting the estimated probability of default/death. We will consider the simple random walk with drift (where $\sigma_t = \sigma$) and the case where $\sigma_t$ follows a GARCH(1,1) with normal errors because both of them allow for closed-form solutions for the ZPP, see \cite{Fantazzini and Zimin, 2020} for details. We will also consider the case where $\sigma_t$ follows a GARCH(1,1) with student’s $t$ errors as originally proposed in \cite{Fantazzini et al., 2008}, and a GARCH(1,1) with errors following the Generalized Hyperbolic Skew-Student distribution proposed by \cite{Aas and Haff, 2006}, which has one tail with polynomial and one with exponential behavior. More recently, \cite{Ardia et al., 2019} and \cite{Maciel, 2021} found that a two-regime Markov-switching GARCH model showed the best in-sample performance when modelling crypto-coins log-returns, and outperformed standard single-regime GARCH models when forecasting the one-day ahead Value-at-Risk. Therefore, we will also use this model in our empirical analysis to compute the ZPP for the first time using a Markov-Switching model.
3.4 Model Evaluation

The main tool to compare the forecasting performances of models with binary data is the confusion matrix by [Provost and Kohavi, 1998], see Table 1:

| Observed/Predicted | DEAD COINS | ALIVE |
|--------------------|------------|-------|
| DEAD COINS         | a          | b     |
| ALIVE              | c          | d     |

Table 1: Theoretical confusion matrix. Number of: a true positive, b false positive, c false negative, d true negative.

In our specific case, the cells of the confusion matrix have the following meaning: a is the number of correct predictions that a coin is dead, b is the number of incorrect predictions that a coin is dead, c is the number of incorrect predictions that a coin is alive, while d is the number of correct predictions that a coin is alive. The confusion matrix is then used to compute the Area Under the Receiver Operating Characteristic curve (AUC or AUROC) proposed by [Metz, 1978], [Metz and Kronman, 1980], and [Hanley and McNeil, 1982] for all forecasting models. The ROC curve is created by plotting, for any probability cut-off value between 0 and 1, the proportion of correctly predicted dead coins \( a/(a+b) \) on the y-axis, also known as sensitivity or hit rate, and the proportion of alive coins predicted as dead coins \( c/(c+d) \) on the x-axis, also known as false-positive rate or as 1 - specificity, where the latter is \( d/(d+c) \).

The AUC lies between zero and one and the closer it is to one the more accurate the forecasting model is, see [Sammut and Webb, 2011], pp. 869-875, and references therein for more details.

Despite its widespread use, the AUC has also some limitations, as discussed in detail by [Krzanowski and Hand, 2009], p. 108. Therefore, we also employed the Model Confidence Set (MCS) proposed by [Hansen et al., 2011] and extended by [Fantazzini and Maggi, 2015] to binary models, to select the best forecasting models among a set of competing models with a specified confidence level. The MCS procedure picks the best forecasting model and computes the probability that the other models are statistically different from the best one using an evaluation rule based on a loss function that, in the case of binary models, is represented by the [Brier, 1950] score. Briefly, the MCS approach tests at each iteration that all models in the set of forecasting models \( M = M_0 \) have an equal forecasting accuracy using the following null hypothesis for a given confidence level \( 1 - \alpha \),

\[
H_{0,M} = E(d_{ij}) = 0, \quad \forall i, j \in M, \quad vs \quad H_{A,M} = E(d_{ij}) \neq 0
\]

where \( d_{ij} = L_i - L_j \) is the sample loss differential between forecasting models \( i \) and \( j \) and \( L_i \) stands for the loss function of model \( i \) (in our case, the Brier score). If the null hypothesis cannot be rejected, then \( \hat{M}_{1-\alpha} = M \). If the null hypothesis is rejected, an elimination rule is used to remove the worst forecasting...
models from the set $M$. The procedure is repeated until the null hypothesis cannot be rejected, and the final set of models defines the so-called model confidence set $\hat{M}_{1-\alpha}^*$. We will employ the T-max statistic for the equivalence test in the MCS procedure. A brief description of this test is reported below, while we refer to [Hansen et al., 2011], for more details. First, the following t-statistics are computed, 

$$t_i = \frac{\bar{d}_i / \hat{\text{var}}(\bar{d}_i)}{\sqrt{\frac{m}{m-1} \sum_{j \in M} \bar{d}_{ij}}}$$

for $i \in M$, where $\bar{d}_i = m^{-1} \sum_{j \in M} \bar{d}_{ij}$ is the simple loss of the $i$-th model relative to the average losses across models in the set $M$, and $\bar{d}_{ij} = H^{-1} \sum_{h=1}^H d_{ij,h}$ measures the sample loss differential between model $i$ and $j$, and $H$ is the number of forecasts. The T-max statistic is then calculated as $T_{\text{max}} = \max_{i \in M}(t_i)$. This statistic has a non-standard distribution that is estimated using bootstrapping methods with 1000 replications. If the null hypothesis is rejected, one model is eliminated using the following elimination rule: 

$$e_{\text{max},M} = \arg \max_{i \in M} \left( \frac{\bar{d}_i / \hat{\text{var}}(\bar{d}_i)}{\sqrt{\frac{m}{m-1} \sum_{j \in M} \bar{d}_{ij}}} \right).$$

### 3.5 Data

We collected the data examined in this paper using two sources of information:

- **https://coinmarketcap.com**: CoinMarketCap is the main aggregator of crypto-coins market data, and it has been owned by the crypto exchange Binance since April 2020, see [https://crypto.marketswiki.com/index.php?title=CoinMarketCap](https://crypto.marketswiki.com/index.php?title=CoinMarketCap). It provides Open-High-Low-Close price data, volume data, market capitalization, and a wide range of additional information.

- **Google Trends**: the Search Volume Index provided by Google Trends shows how many searches have been done for a keyword or a topic on Google over a specific period and a specific region. See [https://support.google.com/trends/?hl=en](https://support.google.com/trends/?hl=en) for more details.

The dataset consisted of 2003 crypto-coins that were alive or dead (according to different criteria) between January 2014 and May 2020. When collecting coin data, we noticed the presence of coins with short time series and coins with long time series. Therefore, we decided to separate coins with less than 750 observations (*young coins*) from the coins with more than 750 observations (*old coins*): we chose this type of grouping because we used the first set of coins to forecast the 1-day and 30-day ahead probabilities of death, while the second set to forecast the 1-day, 30-day, and 365-day ahead probabilities of death, respectively. The effects of different types of groupings are presented in the robustness checks.

As discussed in detail in section 3.1, we employed three competing criteria to classify a coin as dead or alive:

- the approach proposed by [Feder et al., 2018];

- the approach proposed by [Schmitz and Hoffmann, 2020].
• the professional rule that defines a coin dead if its value drops below 1 cent, and alive if its value rises above 1 cent.

The total number of “dead days”, that is the total number of days when the coins are deemed as “dead” according to the previous criteria is reported in Table 2, both in absolute value and percentages.

| YOUNG COINS | OLD COINS |
|-------------|-----------|
| Feder et al. (2018) | Simplified Feder et al. (2018) | 1 cent |
| N. of dead days | % | N. of dead days | % | N. of dead days | % |
| 53169 | 9.89 | 128163 | 23.84 | 310707 | 57.79 |
| 114790 | 11.63 | 428288 | 43.39 | 379226 | 38.42 |

Table 2: Number of dead days (in absolute value and %) for different criteria used to classify a coin as dead or alive.

As expected, the [Feder et al., 2018] approach is the most restrictive with fewer identified dead coins, while the professional rule that defines a coin dead if its value drops below 1 cent is laxer, allowing for a much larger number of dead coins. The simplified [Feder et al., 2018] approach proposed by [Schmitz and Hoffmann, 2020] stays in the middle between the previous two approaches in the case of young coins, whereas it is the least restrictive in the case of old coins\(^\text{13}\).

The total number of coins available each day, and the total number of dead coins each day computed using the previous three criteria and the price and volume data from \url{https://coinmarketcap.com} are reported in Figure 1. The [Feder et al., 2018] approach appears to be more stable than the other two methods, which show much more volatile numbers, instead.

The dataset of young coins ranges between August 2015 and May 2020, while the dataset of old coins between January 2014 and May 2020. Following [Fantazzini and Zimin, 2020], in the case of young coins, we used the lagged average monthly trading volume and the lagged average monthly search volume index provided by Google Trends as regressors for the logit, probit, cauchit, and random forest models. We computed direct forecasts, so we used the 1-day lagged regressors to forecast the 1-day ahead probability of death, while the 30-day lagged regressors to forecast the 30-day ahead probability of death. In the case of old coins, we also added the lagged average yearly trading volume and the lagged average yearly search volume index, and we used the 365-day lagged regressors to forecast the 365-day ahead probability of death.

The first initialization sample used for the estimation of credit scoring and ML models was August 2015 - December 2018 for the young coins, and January 2014 - December 2015 for the old coins. These

\(^{13}\)For ease of reference, we will refer to the [Feder et al., 2018] approach as “restrictive”, to the simplified [Feder et al., 2018] approach as “simple”, while to the professional rule as “1 cent”.

13
time samples were chosen so that the first estimation windows had approximately 100,000 observations\textsuperscript{14}. In simple terms, all coin data were pooled together up to time \( t \) (for example), and the credit scoring and ML models were then fitted to this dataset and the required forecasted probabilities of deaths were computed. After that, the time window was increased by 1 day, and the previous procedure was repeated.

A schematic example of a pooled coin dataset used for credit scoring and ML models is reported in Table

\textsuperscript{14}The experience of the author (both in the academia and in the professional field) with credit risk management for SMEs and with potentially noisy and fraudulent data suggested a minimum dataset of 50,000-100,000 data to have robust estimates.
Table 3: Schematic example of a pooled coin dataset used for credit scoring and ML models.

To deal with potential structural breaks, we considered two types of estimation windows: a rolling fixed window of 100,000 observations and the traditional expanding window.

Time series models using the ZPP were instead estimated separately for each coin. Given that the time series of historical market prices were relatively short (particularly for young coins), we employed only an expanding window scheme with the first estimation sample consisting of 30 observations.\textsuperscript{15}

4 Results

We computed the probability of death for the following two sets of coins:

- 1165 young coins for a total of 537693 observations, whose names are reported in Tables 10-12 in the Appendix. We used this set of coins to forecast the 1-day and 30-day ahead probabilities of death.

- 838 old coins for a total of 987018 observations, whose names are reported in Table 13-14 in the Appendix. We used this set of coins to forecast the 1-day, 30-day, and 365-day ahead probabilities of death.

For the sake of space and interest, given the very large dataset at our disposal, we focused exclusively on out-of-sample forecasting, whereas the in-sample analysis dealing with the models’ residuals was not

\textsuperscript{15}We remark that the datasets used for the estimation of credit scoring, ML models and time series-based models were different, so there were dates for which forecasts from all models were not available. This situation had no impact on individual metrics such as the AUC, but it affected the computation of the Model Confidence Set using the Brier score: in the latter case, we used only dates where forecasts from all models were available.
We computed direct forecasts for the credit scoring and ML models so, at a given time \( t \), we estimated these models as many times as the number of forecast horizons and with regressors lagged as many days as the length of the forecast horizons (1-day lagged regressors to forecast the 1-day ahead probability of death, and so on). Instead, the time series models using the ZPP were estimated only once, and the probabilities of deaths for different forecast horizons were computed using recursive forecasts\(^{17}\).

The AUC scores, the Brier scores, the models included in the Model Confidence Set (MCS), and how many times (in \%) the models did not reach numerical convergence, across the three competing criteria to classify a coin as dead or alive, are reported in Table 4 for the young coins, and in Table 5 for the old coins.

The forecasting metrics for the young coins show that the cauchit model with a fixed estimation window of 100000 observations is generally the best model for all forecast horizons considered and across most criteria to classify a coin as dead or alive. This result confirms the simulation evidence reported in [Gündüz and Fokoué, 2017] who showed that the cauchit is the model of choice under a high level of sparseness of the input space: this is definitely the case for the dataset of young coins, whose trading volumes and Google searches are mostly very low and close to zero. However, we remark that the ZPP computed using a MS-GARCH(1,1) model is the best model when using the professional rule that defines a coin dead if its value drops below 1 cent, thus indirectly confirming the good empirical performances reported in [Ardia et al., 2019] and [Maciel, 2021]. Similarly, according to the AUCs, the ZPP computed using the simple random walk provides good forecasts across all horizons and classifying criteria, which is in line with all past literature dealing with the ZPP.

In the case of old coins, the random forests model with an expanding estimation window is the best model for forecasting the probability of death up to 30 days ahead. Instead, credit scoring models and the ZPP models computed with the random walk and the MS-GARCH(1,1) are the best for the 365-day ahead horizon, according to loss functions and AUCs, respectively. The latter horizon is arguably the most important for credit risk management purposes because this is the time interval that is usually considered by national rules and international agreements, such as the Basel 2 and Basel 3 agreements.

---

\(^{16}\)The author want to thank three anonymous professionals working in the crypto industry for pointing his work in this direction.

\(^{17}\)The development of ZPP models allowing for direct forecasts is left as an avenue for further research.
| Models                              | AUC (restrictive) | AUC (simple) | AUC (1 cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1 cent) | MCS (restrictive) | MCS (simple) | MCS (1 cent) | % not converged |
|------------------------------------|------------------|--------------|--------------|---------------------------|---------------------|---------------------|------------------|-------------|--------------|----------------|
| Logit (expanding window)           | 0.79             | 0.73         | 0.60         | 0.089                     | 0.182               | 0.248               | not included     | not included | not included | 0.00           |
| Probit (expanding window)          | 0.75             | 0.70         | 0.59         | 0.091                     | 0.186               | 0.240               | not included     | not included | not included | 0.00           |
| Cauchit (expanding window)         | 0.86             | 0.80         | 0.64         | 0.077                     | 0.161               | **0.233**           | not included     | not included | included     | 0.00           |
| Random Forest (expanding window)   | 0.78             | 0.76         | 0.72         | 0.080                     | 0.154               | 0.240               | not included     | not included | included     | 0.00           |
| Logit (fixed window)               | 0.84             | 0.77         | 0.58         | 0.081                     | 0.170               | 0.250               | not included     | not included | not included | 0.00           |
| Probit (fixed window)              | 0.83             | 0.74         | 0.58         | 0.083                     | 0.175               | 0.250               | not included     | not included | not included | 0.00           |
| Cauchit (fixed window)             | **0.86**         | **0.80**     |**0.64**      | **0.077**                 | **0.157**           | 0.241               | INCLUDED         | not included | not included | 0.00           |
| Random Forest (fixed window)       | 0.74             | 0.73         | 0.62         | 0.089                     | 0.180               | 0.291               | not included     | not included | not included | 0.00           |
| ZPP - Random walk                  | 0.79             | 0.75         | 0.72         | 0.152                     | 0.199               | 0.384               | not included     | not included | not included | 0.00           |
| ZPP - Normal GARCH(1,1)            | 0.74             | 0.69         | 0.65         | 0.167                     | 0.248               | 0.512               | not included     | not included | not included | 1.70           |
| ZPP - Student’s t GARCH(1,1)       | 0.60             | 0.57         | 0.66         | 0.098                     | 0.244               | 0.532               | not included     | not included | not included | 0.90           |
| ZPP - GH Skew-Student GARCH(1,1)   | 0.62             | 0.59         | 0.44         | 0.099                     | 0.250               | 0.540               | not included     | not included | included     | 43.17          |
| ZPP - MSGARCH(1,1)                 | 0.73             | 0.70         | 0.83         | 0.101                     | 0.241               | 0.469               | not included     | not included | not included | 0.81           |

Table 4: Young coins: AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), models included in the MCS, and numerical convergence failures in percentage across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent”.
## Table 5: Old coins: AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), models included in the MCS, and numerical convergence failures in percentage across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent”.

| Models                      | AUC (restrictive) | AUC (simple) | AUC (1 cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1 cent) | MCS (restrictive) | MCS (simple) | MCS (1 cent) | % not converged |
|-----------------------------|-------------------|--------------|--------------|---------------------------|----------------------|---------------------|-------------------|--------------|--------------|----------------|
| Logit (expanding window)    | 0.74              | 0.71         | 0.68         | 0.104                     | 0.245                | 0.194               | not included      | not included  | not included | 0.00           |
| Probit (expanding window)   | 0.76              | 0.74         | 0.74         | 0.104                     | 0.240                | 0.197               | not included      | not included  | not included | 0.00           |
| Cauchit (expanding window)  | 0.76              | 0.85         | 0.76         | 0.104                     | 0.183                | 0.193               | not included      | not included  | not included | 0.00           |
| Random Forest (expanding window) | 0.78          | 0.84         | 0.77         | 0.087                     | 0.191                | 0.167               | not included      | not included  | not included | 0.00           |
| Logit (fixed window)        | 0.77              | 0.75         | 0.75         | 0.103                     | 0.244                | 0.196               | not included      | not included  | not included | 0.00           |
| Probit (fixed window)       | 0.76              | 0.74         | 0.74         | 0.106                     | 0.228                | 0.202               | not included      | not included  | not included | 0.00           |
| Cauchit (fixed window)      | 0.77              | 0.85         | 0.76         | 0.104                     | 0.183                | 0.193               | not included      | not included  | not included | 0.00           |
| Random Forest (fixed window) | 0.78              | 0.75         | 0.71         | 0.103                     | 0.257                | 0.216               | not included      | not included  | not included | 0.00           |
| ZPP - Normal GARCH(1,1)     | 0.64              | 0.59         | 0.64         | 0.125                     | 0.402                | 0.243               | not included      | not included  | not included | 1.22           |
| ZPP - Student’s GARCH(1,1)  | 0.57              | 0.54         | 0.63         | 0.117                     | 0.387                | 0.248               | not included      | not included  | not included | 1.92           |
| ZPP - GH Skew-Student GARCH(1,1) | 0.57         | 0.55         | 0.42         | 0.120                     | 0.396                | 0.253               | not included      | not included  | not included | 42.70          |
| ZPP - MSGARCH(1,1)          | 0.69              | 0.68         | 0.70         | 0.111                     | 0.374                | 0.229               | not included      | not included  | not included | 0.67           |
In general, our empirical evidence shows that ZPP-based models tend to show better AUCs for long-term forecasts of the probability of death, whereas credit scoring and ML models have better loss functions. This result was expected because the latter models tend to provide smoothed forecasts by construction, while this is not the case for time series-based models. An important advantage of credit scoring and ML models is the greater ease of estimation than the other models. The ZPP computed with the random walk model share the same numerical efficiency, whereas the GARCH(1,1) with errors following the Generalized Hyperbolic Skew-Student distribution had (by far) the worst numerical performance across all datasets: this was not a surprise given that the high complexity of this model is poorly suited for (extremely) noisy data such as crypto coins data.

Given that ZPP-based models seem to better distinguish between future dead and alive coins, while credit scoring and ML models provide smaller loss functions, this evidence strongly suggests the possibility of forecasting gains using forecast combinations methods. We leave this topic as an avenue for future research.

The intuition behind these results is that the additional information provided by trading volumes and Google searches does indeed help to improve the forecasting of the probabilities of deaths, particularly for short-term horizons. We also tried to add these regressors to time-series-based models, but the estimation of the models turned out to be either poor or not viable due to the short time series available for estimation, and for this reason, we did not consider such models\textsuperscript{18}. It is well-known since the work by [Fiorentini et al., 1996] that the estimation of GARCH models is complex and requires large samples. Moreover, the large simulation studies of GARCH processes in [Hwang and Valls Pereira, 2006], [Fantazzini, 2009], and [Bianchi et al., 2011] showed that a sample of at least 250-500 observations is needed to have good model estimates and, in case of complex data generating processes, even larger samples are required.

5 Robustness checks

We wanted to verify that our previous results also held with different data samples. Therefore, we performed a series of robustness checks considering the models’ forecasting performances before and after the burst of the bitcoin bubble at the end of 2017, and when separating crypto coins with large market capitalization from coins with small market capitalization.

\textsuperscript{18}We also tried to add these regressors in the mean equation of the simple random walk model, but the results did not change qualitatively (results not reported). This was not a surprise because it is the variance modelling that is the key ingredient when computing the ZPP, see [Fantazzini and Zimin, 2020] -section 4.3- and references therein for more details.
5.1 Forecasting the probability of death before and after the 2017 bubble

There is increasing literature showing that there was a financial bubble in the bitcoin prices in 2016-2017 that burst at the end of 2017, see [Fry, 2018], [Corbet et al., 2018], [Gerlach et al., 2019], and [Xiong et al., 2020]. Besides, there is also a debate on whether the introduction of bitcoin futures in December 2017 crashed the market prices, see [Koehling et al., 2019], [Liu et al., 2020], [Baig et al., 2020], [Jalan et al., 2021], and [Hattori and Ishida, 2021]. [Fantazzini and Kolodin, 2020] used several unit root tests allowing for an endogenous break and found a significant structural break located at the end of 2017, so they fixed a break date on 10 December 2017 that is the day when the first bitcoin futures were introduced on the CBOE.

Following this literature, we divided our dataset into two sub-samples consisting of data before and after 10 December 2017, and we examined the models’ forecasting performances in these two sub-samples. Given the very small number of young coins available before the end of 2017, we only considered old coins for this robustness check (that is coins with at least 750 observations).

The AUC scores, the Brier scores, and the models included in the Model Confidence Set (MCS) across the three competing criteria to classify a coin as dead or alive, are reported in Table 6 for the sub-sample ending on 10 December 2017, and in Table 7 for the sub-sample starting after that date.

Tables 6 and 7 do not highlight any major differences between the two sub-samples. However, we can notice that the general levels of the AUCs for the 30-day and 365-days forecast horizons slightly decreased in the second sub-sample after the burst of the 2017 bubble. Moreover, in the latter sub-sample, credit scoring models (particularly the cauchit) showed better results compared to the random forest and ZPP models than in the first sub-sample, that is before the bubble burst. Probably, the fall in trading volumes and Google searches after 2017 increased the sparseness of the input space, thus favoring models such as the cauchit, as shown by [Gündüz and Fokoué, 2017] and discussed in the previous pages.
### OLD COINS: 1-DAY AHEAD PROBABILITY OF DEATH (2016 -2017)

| Models                               | AUC (restrictive) | AUC (simple) | AUC (1cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1cent) | MCS (restrictive) | MCS (simple) | MCS (1cent) |
|--------------------------------------|-------------------|--------------|-------------|---------------------------|----------------------|---------------------|--------------------|---------------|-------------|
| Logit (expanding window)             | 0.76              | 0.74         | 0.76        | 0.189                     | 0.209                | 0.299               | INCLUDED          |              |             |
| Probit (expanding window)            | 0.71              | 0.69         | 0.76        | 0.103                     | 0.215                | 0.238               | not included      |              |             |
| Cauchit (expanding window)           | 0.70              | 0.81         | 0.079       | 0.142                     | 0.195                | 0.219               | not included      |              |             |
| Random Forest (expanding window)     | 0.97              | 0.96         | 0.025       | 0.052                     | 0.066                | INCLUDED            | INCLUDED          |              | INCLUDED    |
| Legit (fixed window)                 | 0.77              | 0.81         | 0.086       | 0.147                     | 0.198                | 0.219               | not included      |              |             |
| Probit (fixed window)                | 0.71              | 0.69         | 0.79        | 0.219                     | 0.204                | 0.219               | not included      |              |             |
| Cauchit (fixed window)               | 0.81              | 0.84         | 0.079       | 0.137                     | 0.184                | 0.219               | not included      |              |             |
| Random Forest (fixed window)         | 0.93              | 0.92         | 0.039       | 0.083                     | 0.117                | 0.154               | not included      |              |             |
| ZPP - Random walk                    | 0.81              | 0.76         | 0.105       | 0.202                     | 0.292                | 0.312               | not included      |              |             |
| ZPP - Normal GARCH(1,1)              | 0.60              | 0.60         | 0.65        | 0.249                     | 0.307                | 0.312               | not included      |              |             |
| ZPP - Student’s GARCH(1,1)           | 0.56              | 0.51         | 0.37        | 0.236                     | 0.312                | 0.312               | not included      |              |             |
| ZPP - GH Skew-Student GARCH(1,1)     | 0.55              | 0.51         | 0.43        | 0.240                     | 0.315                | 0.312               | not included      |              |             |
| ZPP - MSGARCH(1,1)                   | 0.71              | 0.71         | 0.83        | 0.189                     | 0.252                | 0.289               | not included      |              |             |

Table 6: Old coins: years 2016 - 2017. AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), and models included in the MCS across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent”. 
| Models                        | AUC (restrictive) | AUC (simple) | AUC (1cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1cent) | MCS (restrictive) | MCS (simple) | MCS (1cent) |
|-----------------------------|-------------------|--------------|-------------|---------------------------|----------------------|---------------------|------------------|--------------|-------------|
| Logit (expanding window)    | 0.76              | 0.74         | 0.66        | 0.130                     | 0.213                | 0.184               | not included     | not included | not included |
| Probit (expanding window)   | 0.76              | 0.74         | 0.66        | 0.120                     | 0.247                | 0.187               | not included     | not included | not included |
| Cauchit (expanding window)  | 0.78              | 0.77         | 0.67        | 0.110                     | 0.173                | 0.177               | not included     | not included | not included |
| Random Forest (expanding window) | 0.96            | 0.97         | 0.95        | 0.037                     | 0.068                | 0.070               | INCLUDED         | INCLUDED     | INCLUDED     |
| Logit (fixed window)        | 0.79              | 0.74         | 0.73        | 0.108                     | 0.244                | 0.195               | not included     | not included | not included |
| Probit (fixed window)       | 0.79              | 0.76         | 0.72        | 0.108                     | 0.230                | 0.202               | not included     | not included | not included |
| Cauchit (fixed window)      | 0.79              | 0.86         | 0.73        | 0.111                     | 0.195                | 0.196               | not included     | not included | not included |
| Random Forest (fixed window) | 0.74              | 0.82         | 0.72        | 0.100                     | 0.220                | 0.181               | not included     | not included | not included |
| ZPP - Random walk           | 0.76              | 0.73         | 0.75        | 0.203                     | 0.272                | 0.196               | not included     | not included | not included |
| ZPP - Normal GARCH(1,1)     | 0.64              | 0.59         | 0.64        | 0.127                     | 0.442                | 0.227               | not included     | not included | not included |
| ZPP - Student's t GARCH(1,1) | 0.57              | 0.53         | 0.63        | 0.122                     | 0.426                | 0.231               | not included     | not included | not included |
| ZPP - GH Skew-Student GARCH(1,1) | 0.57              | 0.54         | 0.42        | 0.125                     | 0.437                | 0.234               | not included     | not included | not included |

Table 7: Old coins: years 2018 - 2020. AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), and models included in the MCS across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent.”
5.2 Large cap and small cap: does it matter?

In the baseline case, we separated our coins data based on the length of their time series for forecasting purposes. Moreover, before starting our analysis, we tried different clustering methods to group coins with similar attributes, and most methods proposed groupings quite close to our simple baseline approach\(^{19}\). However, we also noticed that some methods separated the 50-100 coins with the largest market capitalizations from all others. Therefore, we separated the 100 crypto coins with the largest market capitalization from all other coins with a smaller market capitalization, and we examined how the models’ forecasting performances changed.

The AUC scores, the Brier scores, and the models included in the Model Confidence Set (MCS) across the three competing criteria to classify a coin as dead or alive, are reported in Table 8 for the 100 coins with the largest market capitalization, and in Table 9 for all other coins.

Tables 8 and 9 show that the separation of coins based on their market capitalization did not produce any major changes compared to the baseline case. However, there are some differences: in the case of big-cap coins, the random forests model remained the best model only for 1-day ahead forecasts, whereas the cauchit was the best model for both the 30-day and 365-day ahead forecast horizons. A similar picture also emerged for small-cap coins, where credit scoring models and the ZPP computed with the MS-GARCH(1,1) were the best models for the 30-day and 365-day ahead forecast horizons. Interestingly, the success of credit scoring and ZPP-based models for the long-term forecasts of the probability of death of small-cap coins are qualitatively similar to the evidence reported by [Fantazzini and Zimin, 2020] who used only 42 coins (most of them small-cap).

\(^{19}\)See [Romesburg, 2004] and [Everitt, 2011] for an introduction to cluster analysis at the textbook level.
| Models                  | AUC (restrictive) | AUC (simple) | AUC (1cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1cent) | MCS (restrictive) | MCS (simple) | MCS (1cent) |
|------------------------|------------------|--------------|-------------|---------------------------|----------------------|---------------------|--------------------|--------------|--------------|
| Logit (expanding window) | 0.86             | 0.84         | 0.76        | 0.004                     | 0.004                | 0.004               | INCLUDED          | INCLUDED     | INCLUDED     |
| Probit (expanding window) | 0.86             | 0.86         | 0.75        | 0.020                     | 0.101                | 0.086               | not included       | not included  | not included  |
| Cauhitz (expanding window) | 0.90             | 0.90         | 0.74        | 0.007                     | 0.072                | 0.093               | INCLUDED          | not included  | not included  |
| Random Forest (expanding window) | 0.96             | 0.97         | 0.96        | 0.003                     | 0.027                | 0.032               | INCLUDED          | INCLUDED     | INCLUDED     |
| Legit (fixed window) | 0.82             | 0.66         | 0.66        | 0.006                     | 0.084                | 0.106               | INCLUDED          | not included  | not included  |
| Probit (fixed window) | 0.83             | 0.66         | 0.63        | 0.010                     | 0.087                | 0.106               | not included       | not included  | not included  |
| Cauhitz (fixed window) | 0.89             | 0.85         | 0.75        | 0.005                     | 0.079                | 0.301               | INCLUDED          | not included  | not included  |
| Random Forest (fixed window) | 0.66             | 0.63         | 0.62        | 0.006                     | 0.093                | 0.106               | INCLUDED          | not included  | not included  |
| ZPP - Random walk | 0.83             | 0.83         | 0.89        | 0.036                     | 0.079                | 0.126               | not included       | not included  | not included  |
| ZPP - Normal GARCH(1,1) | 0.64             | 0.54         | 0.60        | 0.006                     | 0.100                | 0.097               | INCLUDED          | not included  | not included  |
| ZPP - Student's t GARCH(1,1) | 0.73             | 0.56         | 0.29        | 0.006                     | 0.097                | 0.098               | INCLUDED          | not included  | not included  |
| ZPP - GH Skew-Student GARCH(1,1) | 0.65            | 0.58         | 0.39        | 0.006                     | 0.098                | 0.098               | INCLUDED          | not included  | not included  |
| ZPP - MSGARCH(1,1) | 0.76             | 0.69         | 0.62        | 0.006                     | 0.093                | 0.091               | INCLUDED          | not included  | not included  |

Table 8: Big cap coins: AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), and models included in the MCS across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent”.
| Models | AUC (restrictive) | AUC (simple) | AUC (1cent) | Brier Score (restrictive) | Brier Score (simple) | Brier Score (1cent) | MCS (restrictive) | MCS (simple) | MCS (1cent) |
|--------|-----------------|--------------|-------------|---------------------------|----------------------|---------------------|------------------|---------------|-------------|
| Logit (expanding window) | 0.74 | 0.75 | 0.67 | 0.111 | 0.224 | 0.219 | not included | not included | not included |
| Probit (expanding window) | 0.72 | 0.73 | 0.66 | 0.117 | 0.234 | 0.222 | not included | not included | not included |
| Cauchit (expanding window) | 0.79 | 0.84 | 0.72 | 0.103 | 0.173 | 0.207 | not included | not included | not included |
| Random Forest (expanding window) | 0.90 | 0.92 | 0.89 | 0.101 | 0.132 | 0.132 | INCLUDED | INCLUDED | INCLUDED |
| Logit (fixed window) | 0.77 | 0.75 | 0.72 | 0.105 | 0.218 | 0.223 | not included | not included | not included |
| Probit (fixed window) | 0.76 | 0.74 | 0.71 | 0.107 | 0.222 | 0.228 | not included | not included | not included |
| Cauchit (fixed window) | 0.78 | 0.82 | 0.74 | 0.104 | 0.183 | 0.218 | not included | not included | not included |
| Random Forest (fixed window) | 0.76 | 0.82 | 0.76 | 0.096 | 0.196 | 0.216 | not included | not included | not included |
| ZPP - Random walk | 0.76 | 0.74 | 0.69 | 0.185 | 0.253 | 0.283 | not included | not included | not included |
| ZPP - Normal GARCH(1,1) | 0.65 | 0.59 | 0.64 | 0.130 | 0.375 | 0.351 | not included | not included | not included |
| ZPP - Student's t GARCH(1,1) | 0.58 | 0.54 | 0.65 | 0.120 | 0.363 | 0.361 | not included | not included | not included |
| ZPP - GH Skew-Student GARCH(1,1) | 0.58 | 0.56 | 0.41 | 0.123 | 0.372 | 0.361 | not included | not included | not included |
| ZPP - MSGARCH(1,1) | 0.69 | 0.67 | 0.73 | 0.117 | 0.219 | 0.219 | not included | not included | not included |

Table 9: Small cap coins: AUC scores (highest values are in bold fonts), Brier scores (smallest values are in bold fonts), and models included in the MCS across three competing criteria to classify a coin as dead or alive. [Feder et al., 2018] approach = “restrictive”; simplified [Feder et al., 2018] approach = “simple”; professional rule = “1 cent”.

6 Conclusions

This paper examined a set of over two thousand crypto-coins observed between 2015 and 2020, to estimate their credit risk by computing their probability of death using different definitions of dead coins, and different forecasting horizons.

To reach this aim, we first employed a set of models to forecast the probability of death including credit scoring models, machine learning models, and time-series methods based on the Zero Price Probability (ZPP) model, which is a methodology to compute the probabilities of default using only market prices. Secondly, we performed a forecasting exercise using a unique set of 2003 crypto coins that were active from the beginning of 2014 till the end of May of 2020. Our results showed that the choice of the coin death definition significantly affected the set of the best forecasting models to compute the probability of death. However, this choice was not critical, and the best models turned out to be the same in most cases. In general, we found that the cauchit and the ZPP based on the random walk or the MS-GARCH(1,1) were the best models for newly established coins, whereas credit scoring models and machine learning methods using lagged trading volumes and online searches were better choices for older coins.

Finally, we performed a set of robustness checks to verify that our results also held with different data samples. To achieve this aim, we considered the models' forecasting performances before and after the burst of the bitcoin bubble at the end of 2017, and when we separated crypto coins with large market capitalization from coins with small market capitalization. The two robustness checks did not produce any major changes compared to the baseline case.

The general recommendation for investors that emerged from our analysis is to use the cauchit model when dealing with coins with a short time series and/or with trading volumes and Google searches close to zero. In the case of a large information set and the main interest is on short-term forecasting, the random forests model is definitely the model of choice, whereas the ZPP-based models using the simple random walk or the MS-GARCH(1,1) are to be preferred in case of long-term forecasts up to 1-year ahead.

Another implication of the findings of our work is the need to have more transparency and better reporting about the credit risk of crypto assets. Despite the large losses incurred by investors in the last years, the lack of focus on risk management practices is somewhat astonishing. One of the best practices that this work clearly suggest is for crypto exchanges to publish daily the estimated probability of death for the traded crypto assets, using one of the models discussed in this paper, or the simple average of the estimates provided by several models. The reported probabilities of death would warn investors about the risk of investing in crypto assets, thus helping them making more considered investment decisions.

We should note that our empirical analysis highlighted that the major drawback of the ZPPs computed...
using GARCH models is the need to have time series long enough to have decent parameter estimates. This problem makes them unsuitable for newly-established coins. Moreover, the extreme volatility of crypto-coins markets and the frequent presence of structural breaks make things worse. Therefore, it was not a surprise that the ZPPs calculated using the simple random walk or the Markov-Switching GARCH(1,1) model were the best in this class of models. The retrieval of high-frequency data and the use of Bayesian methods to solve these computational issues are left as avenues for future research.

Another possibility of future work will be to explore the feasibility of forecast combinations methods. Given that ZPP-based models seem to better distinguish between future dead and alive coins, while credit scoring and ML models provide smaller loss functions, our empirical evidence suggests the possibility of forecasting gains using combinations methods. This is why this extension could be an interesting issue for future research.
References

Kjersti Aas and Ingrid Hobæk Haff. The generalized hyperbolic skew student's t-distribution. *Journal of financial econometrics*, 4(2):275–309, 2006.

Andreas Antonopoulos. *Mastering Bitcoin: unlocking digital cryptocurrencies*. " O'Reilly Media, Inc."., 2014.

David Ardia, Keven Bluteau, and Maxime Rüede. Regime changes in Bitcoin GARCH volatility dynamics. *Finance Research Letters*, 29:266–271, 2019.

Bart Baesens and Tony Van Gestel. *Credit risk management: Basic concepts*. Oxford University Press, 2009.

Ahmed S Baig, Omair Haroon, and Nasim Sabah. Price clustering after the introduction of bitcoin futures. *Applied Finance Letters*, 9:36–42, 2020.

Flavio Barboza, Herbert Kimura, and Edward Altman. Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83:405–417, 2017.

Basel Committee on Banking Supervision. Findings on the Interaction of Market and Credit Risk. Technical report, BCBS Working Papers, n. 16, 2009.

Carluccio Bianchi, Maria Elena De Giuli, Dean Fantazzini, and Mario Maggi. Small sample properties of copula-GARCH modelling: a Monte Carlo study. *Applied financial economics*, 21(21):1587–1597, 2011.

Glenn Brier. Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3, 1950.

Chris Brummer. *Cryptoassets: legal, regulatory, and monetary perspectives*. Oxford University Press, 2019.

Shaen Corbet, Brian Lucey, and Larisa Yarovaya. Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters*, 26:81–88, 2018.

Luciana Dalla Valle, Maria Elena De Giuli, Claudia Tarantola, and Claudio Manelli. Default probability estimation via pair copula constructions. *European Journal of Operational Research*, 249(1):298–311, 2016.

Marcos Lopez De Prado. *Advances in financial machine learning*. John Wiley & Sons, 2018.

Matthew F Dixon, Igor Halperin, and Paul Bilokon. *Machine Learning in Finance*. Springer, 2020.

Bri...
Dean Fantazzini and Silvia Figini. Random survival forests models for sme credit risk measurement. *Methodology and Computing in Applied Probability*, 11(1):29–45, 2009.

Dean Fantazzini and Nikita Kolodin. Does the hashrate affect the bitcoin price? *Journal of Risk and Financial Management*, 13(11):263, 2020.

Dean Fantazzini and Mario Maggi. Proposed coal power plants and coal-to-liquids plants in the us: Which ones survive and why? *Energy Strategy Reviews*, 7:9–17, 2015.

Dean Fantazzini and Stephan Zimin. A multivariate approach for the simultaneous modelling of market risk and credit risk for cryptocurrencies. *Journal of Industrial and Business Economics*, 47(1):19–69, 2020.

Dean Fantazzini, Maria Elena De Giuli, and Mario Alessandro Maggi. A new approach for firm value and default probability estimation beyond merton models. *Computational Economics*, 31(2):161–180, 2008.

Amir Feder, Neil Gandal, Hamrick, Tyler Moore, and Marie Vasék. The rise and fall of cryptocurrencies. In *17th Workshop on the Economics of Information Security (WEIS)*, 2018. URL https://tylermoore.utulsa.edu/weis18.pdf.

Gabriele Fiorentini, Giorgio Calzolari, and Lorenzo Panattoni. Analytic derivatives and the computation of GARCH estimates. *Journal of applied econometrics*, 11(4):399–417, 1996.

John Fry. Booms, busts and heavy-tails: The story of bitcoin and cryptocurrency markets? *Economics Letters*, 171:225–229, 2018.

Ana-Maria Fuertes and Elena Kalotychou. Early warning systems for sovereign debt crises: The role of heterogeneity. *Computational Statistics and Data Analysis*, 51(2):1420–1441, 2006.

Neil Gandal, JT Hamrick, Tyler Moore, and Tali Oberman. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics*, 95:86–96, 2018.

Neil Gandal, JT Hamrick, Tyler Moore, and Marie Vasek. The rise and fall of cryptocurrency coins and tokens. *Decisions in Economics and Finance*, pages 1–34, 2021.

Jan-Christian Gerlach, Guilherme Demos, and Didier Sornette. Dissection of bitcoin’s multiscale bubble history from January 2012 to February 2018. *Royal Society open science*, 6(7):180643, 2019.

Paolo Giudici and Silvia Figini. *Applied data mining for business and industry*. Wiley Online Library, 2009.

John Griffin and Amin Shams. Is Bitcoin really untethered? *The Journal of Finance*, 75(4):1913–1964, 2020.

Klaus Gurobi and Niranjan Sapkota. Predicting cryptocurrency defaults. *Applied Economics*, 52(46):5060–5076, 2020.

Necla Gündüz and Ernest Fokoué. On the predictive properties of binary link functions. *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 66(1):1–18, 2017.

JT Hamrick, Farhang Rouhi, Arghya Mukherjee, Amir Feder, Neil Gandal, Tyler Moore, and Marie Vasek. An examination of the cryptocurrency pump-and-dump ecosystem. *Information Processing & Management*, 58(4):102506, 2021.

James Hanley and Barbara McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1):29–36, 1982.

Peter Hansen, Asger Lunde, and James Nason. The model confidence set. *Econometrica*, 79(2):453–497, 2011.
Philipp Hartmann. Interaction of market and credit risk, 2010.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. The elements of statistical learning: data mining, inference, and prediction. Springer, 2009.

Takahiro Hattori and Ryo Ishida. Did the introduction of bitcoin futures crash the bitcoin market at the end of 2017? The North American Journal of Economics and Finance, 56:101322, 2021.

Soosung Hwang and Pedro Valls Pereira. Small sample properties of GARCH estimates and persistence. The European Journal of Finance, 12(6-7):473–494, 2006.

Rob Hyndman and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2018.

Akanksha Jalan, Roman Matkovskyy, and Andrew Urquhart. What effect did the introduction of bitcoin futures have on the bitcoin spot market? The European Journal of Finance, pages 1–31, 2021.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. An introduction to statistical learning, volume 112. Springer, 2013.

Jiabao Jing, Wenwen Yan, and Xiaomei Deng. A hybrid model to estimate corporate default probabilities in China based on zero-price probability model and long short-term memory. Applied Economics Letters, 28(5):413–420, 2021.

Ciby Joseph. Advanced credit risk analysis and management. John Wiley & Sons, 2013.

Gerrit Köchling, Janis Müller, and Peter N Posch. Does the introduction of futures improve the efficiency of bitcoin? Finance Research Letters, 30:367–370, 2019.

Roger Koenker and Jungmo Yoon. Parametric links for binary choice models: A fisherian–bayesian colloquy. Journal of Econometrics, 152(2):120–130, 2009.

Wojtek Krzanowski and David Hand. ROC curves for continuous data. Crc Press, 2009.

Lili Li, Jun Yang, and Xin Zou. A study of credit risk of Chinese listed companies: ZPP versus KMV. Applied Economics, 48(29):2697–2710, 2016.

Ruozhou Liu, Shanfeng Wan, Zili Zhang, and Xuejun Zhao. Is the introduction of futures responsible for the crash of bitcoin? Finance Research Letters, 34:101259, 2020.

Leandro Macel. Cryptocurrencies value-at-risk and expected shortfall: Do regime-switching volatility models improve forecasting? International Journal of Finance & Economics, 26(3):4840–4855, 2021.

Peter McCullagh and John A. Nelder. Generalized Linear Model. Chapman Hall, 1989.

Charles Metz. Basic principles of ROC analysis. In Seminars in nuclear medicine, volume 8, pages 283–298. Elsevier, 1978.

Charles Metz and Helen Kronman. Statistical significance tests for binormal ROC curves. Journal of Mathematical Psychology, 22(3):218–243, 1980.

Mirko Moscatelli, Fabio Parlapiano, Simone Narizzano, and Gianluca Viggiano. Corporate default forecasting with machine learning. Expert Systems with Applications, 161:113567, 2020.

Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton University Press, 2016.
Foster Provost and R Kohavi. Glossary of terms. *Journal of Machine Learning*, 30(2-3):271–274, 1998.

Arnulfo Rodriguez and Pedro N Rodriguez. Understanding and predicting sovereign debt rescheduling: a comparison of the areas under receiver operating characteristic curves. *Journal of Forecasting*, 25(7):459–479, 2006.

Charles Romesburg, *Cluster analysis for researchers*. Lulu. com, 2004.

Claude Sammut and Geoffrey Webb, *Encyclopedia of machine learning*. Springer, 2011.

Fabian Schar and Aleksander Berentsen. *Bitcoin, Blockchain, and Cryptoassets: A Comprehensive Introduction*. MIT Press, 2020.

Tim Schmitz and Ingo Hoffmann. Re-evaluating cryptocurrencies’ contribution to portfolio diversification—a portfolio analysis with special focus on german investors. *Available at SSRN 3625458*, 2020.

Sid. How Peng Coin Will Surge 8-12x These Coming Weeks. *Medium*, July 8, 2018. URL https://medium.com/@sidbicious123/how-peng-coin-will-surge-8-12x-these-coming-weeks-4026831b31c1.

Sandeep Soni. RIP cryptocurrencies: Number of ‘dead’ coins up 35% over last year; tally nears 2,000-mark. *Financial Express*, April 3, 2021. URL https://www.financialexpress.com/market/rip-cryptocurrencies-number-of-dead-coins-up-35-over-last-year-tally-nears-2000-mark/.

En-Der Su and Shih-Ming Huang. Comparing firm failure predictions between logit, KMV, and ZPP models: Evidence from Taiwan’s electronics industry. *Asia-Pacific Financial Markets*, 17(3):209–239, 2010.

Wang Chun Wei. The impact of Tether grants on Bitcoin. *Economics Letters*, 171:19–22, 2018.

Jinwu Xiong, Qing Liu, and Lei Zhao. A new method to verify bitcoin bubbles: Based on the production cost. *The North American Journal of Economics and Finance*, 51:101095, 2020.
Appendix: lists of young and old coins

| #  | Token Name | Table 10: Names of the 1165 young coins: coins 1-400. |
|----|------------|------------------------------------------------------|
| 1  | BitCoin X  | 101 HotPepper Token                                 |
| 2  | Crypto.com Coin | 102 Huobi Token                                     |
| 3  | KuCoin | 103 USD Coin                                       |
| 4  | UNOS SED LEO | 104 Qubitinea Token                                 |
| 5  | USDC | 105 $MK Token                                      |
| 6  | Heco | 106 Ocean Protocol                                   |
| 7  | Huobi | 107 BobCoin Token                                   |
| 8  | VIP Coin | 108 Origin Protocol                                  |
| 9  | BreakTrend | 109 XamarCoin Token                                 |
| 10  | IDO COIN | 110 Project Pat                                    |
| 11  | OKB | 111 Wink                                           |
| 12  | FTX Token | 112 Function X                                      |
| 13  | VeChain | 113 Fetch.ai                                       |
| 14  | Faze Standard | 114 tscoin                                       |
| 15  | Mints/Wh0leCoin | 115 WowToken                                       |
| 16  | PlayFul | 116 Gus                                             |
| 17  | Beldora Hedogaph | 117 Aurora                                          |
| 18  | Skrill | 118 BittuCoin                                      |
| 19  | Largo Coin | 119 SynchroBitCoin                                  |
| 20  | Binance USD | 120 DAI                                            |
| 21  | Hyperion | 121 Eternal Estate                                  |
| 22  | The Make Touch Gold | 122 AgoraCoin                                      |
| 23  | Inlogue Coin | 123 Folgory Coin                                    |
| 24  | ThorCoin | 124 BittKoin                                       |
| 25  | Eternity | 125 DcfChain                                        |
| 26  | Elinarcoin | 126 Uthligate                                       |
| 27  | MINDSOL | 127 Nash Exchange                                   |
| 28  | Dai | 128 NEXT                                            |
| 29  | Baar Chain | 129 Lokibla                                          |
| 30  | HUSD | 130 BigOneToken                                      |
| 31  | Fio Network | 131 WORM Protocol                                   |
| 32  | Velas | 132 BittCoin                                       |
| 33  | Interplanetary Chain Architecture | 133 CONTRADEX                                      |
| 34  | 2B Token | 134 Rocket Pool                                     |
| 35  | GlinkCoin | 135 IBDX                                            |
| 36  | becCoin | 136 Egyrane                                         |
| 37  | Dey | 137 LuckySevenToken                                  |
| 38  | IOT | 138 Jowel                                           |
| 39  | Da@Chain Token | 139 Color Network                                    |
| 40  | Quant | 140 Rumors                                          |
| 41  | Seek-N | 141 Kosmos                                          |
| 42  | Commerce Coin | 142 General Artisanian Currency                    |
| 43  | Nexo Network | 143 Everipedia                                       |
| 44  | CryptoNetwork | 144 Cryptopool Token                                 |
| 45  | Blockstack | 145 Bitcoin 2                                       |
| 46  | Eneryi | 146 ApolloCurves                                    |
| 47  | Chilir | 147 BOBA                                           |
| 48  | iChain | 148 Cryptomemo.com 100                              |
| 49  | BitPrive | 149 GoChain                                         |
| 50  | ABC Coin | 150 MoneroCoin                                      |
| 51  | Ubit | 151 DAI                                             |
| 52  | NewYork Exchange | 152 Bit-E Token                                     |
| 53  | Bifold | 153 Bifold                                          |
| 54  | EarthToken | 154 Machine Xchange Coin                           |
| 55  | BalckGaming | 155 CVN Chain                                       |
| 56  | ZEON | 156 KNC                                             |
| 57  | Solomonic | 158 Neutrin Dollar                                  |
| 58  | Terra | 159 Wasitx                                         |
| 59  | RSK Infrastructure Framework | 160 Nima Y                                      |
| 60  | cymc | 161 BHPCoin                                         |
| 61  | PAX Gold | 162 FastCoin                                        |
| 62  | BitcoinBD | 163 New coin                                        |
| 63  | Elion | 164 The Force Protocol                              |
| 64  | Bloommed Token | 165 COIN                                            |
| 65  | TRON | 166 ICON                                            |
| 66  | Isole | 167 Etherum Meta                                    |
| 67  | Trustee | 168 Trustee                                         |
| 68  | Crypt0sucks | 169 uSD                                             |
| 69  | STEMI CRL Coin | 170 VelaCoin                                        |
| 70  | APIK | 171 Ami                                            |
| 71  | Tap | 172 Champions                                       |
| 72  | Bander | 173 Raken                                           |
| 73  | Frostcoin | 174 Transcoin                                       |
| 74  | FABRKS | 175 Throatn fase                                     |
| 75  | Bitcoin Treasure | 176 BitForen Token                                 |
| 76  | BIREX Token | 177 Wrapped Bitcoin                                |
| 77  | Theta Pool | 178 ZBG Token                                       |
| 78  | Gatchain Token | 179 Orchid                                          |
| 79  | STASIS EURO | 180 TTS                                            |
| 80  | Kero | 181 LGT Network                                     |
| 81  | BTU Protocol | 182 Microtream                                      |
| 82  | ThunderToken | 183 Contento Coin                                   |
| 83  | Beam | 184 Landlord                                        |
| 84  | Swge | 185 Constellation                                   |
| 85  | Reserve Rights | 186 Ultra                                           |
| 86  | Ignition Fortune | 187 FIBOS                                          |
| 87  | Orbi | 188 DREP                                            |
| 88  | Bagttra Coin Zero | 189 Inferno Hyperion Fund                         |
| 89  | LeiX | 190 COWIN                                           |
| 90  | inX | 191 Standard Tokenization Protocol                   |
| 91  | Electra Coin | 192 Mainchain                                       |
| 92  | USDK | 193 Chronia                                         |
| 93  | Viper Super Zero | 194 ARPA Chain                                     |
| 94  | Heido Pool Token | 195 RepToken                                        |
| 95  | Harmony | 196 Carvy                                          |
| 96  | Poseidon Network | 197 Voke Token                                     |
| 97  | Sandchain | 198 Zeno                                            |
| 98  | B2Ships | 199 Eltoun                                         |
| 99  | Vital | 200 Emutex Token                                    |

Table 10: Names of the 1165 young coins: coins 1-400.
Table 11: Names of the 1165 young coins: coins 401-800.
| Coin                  | Rank | Website/Blockchain          | Notes                  |
|----------------------|------|----------------------------|------------------------|
| Bitcoin              | 1    | Bitcoin.org                 |                        |
| Ethereum             | 2    | Ethereum.org                |                        |
| Litecoin             | 3    | Litecoin.org                |                        |
| Ripple               | 4    | Ripple.org                  |                        |
| Bitcoin Cash         | 5    | bitcoin.org                 |                        |
| Monero               | 6    | monero.org                  |                        |
| Zcash                | 7    | zcash.org                   |                        |
| Dash                 | 8    | dash.org                    |                        |
| Litecoin Cash        | 9    | litecoin.org                |                        |
| Bitcoin SV           | 10   | bsv.org                     |                        |
| Dogecoin             | 11   | dogecoin.org                |                        |
| Ripple               | 12   | xrp.org                     |                        |
| Zcash                | 13   | zcash.org                   |                        |
| Monero               | 14   | monero.org                  |                        |
| Zcash                | 15   | zcash.org                   |                        |
| Bitcoin Cash         | 16   | bitcoin.org                 |                        |
| Monero               | 17   | monero.org                  |                        |
| Zcash                | 18   | zcash.org                   |                        |
| Bitcoin Cash         | 19   | bitcoin.org                 |                        |
| Monero               | 20   | monero.org                  |                        |
| Zcash                | 21   | zcash.org                   |                        |
| Bitcoin Cash         | 22   | bitcoin.org                 |                        |
| Monero               | 23   | monero.org                  |                        |
| Zcash                | 24   | zcash.org                   |                        |
| Bitcoin Cash         | 25   | bitcoin.org                 |                        |
| Monero               | 26   | monero.org                  |                        |
| Zcash                | 27   | zcash.org                   |                        |
| Bitcoin Cash         | 28   | bitcoin.org                 |                        |
| Monero               | 29   | monero.org                  |                        |
| Zcash                | 30   | zcash.org                   |                        |

Table 13: Names of the 838 old coins: coins 1-420.
Table 14: Names of the 838 old coins: coins 401-838.