International Journal on Cybernetics & Informatics (IJCI) Vol.13, No.2, April 2024

THE MATHEMATICS BEHIND CRYPTOCURRENCIES"A STATISTICAL ANALYSIS OF CRYPTOCURRENCIES

Masoud Eshaghinasrabadi

California State University Northridge

ABSTRACT

This article provides a statistical approach to describe the fit of the most popular cryptocurrencies, building off a previous report, "A Statistical Analysis of Cryptocurrencies." We examined Bitcoin, Ethereum, Tether, Binance, Ripple, Cardano, Solana, and Doge coins. To model our cryptocurrencies, we utilized trading prices between 2017 and 2022 in light of historic events, such as the COVID-19 pandemic. Additionally, we performed a correlation analysis to help understand the relationship between the popular cryptos. Here, we report that the candidate distributions we fit to model the currencies needed to be more independent to describe the return of all popular cryptos. This could be due to the need for Correlation between some of these popular cryptos. We found the generalized hyperbolic and the generalized t showed the best performance of the models tested, though these approaches remained limited in their overall fitness. Their performance also varied by cryptocurrency under investigation, with Tether demonstrating the worst fit across all candidate models. Using our fit models, we also predicted the average daily returns for January 1st, 2023, to February 1st, 2023, and generally found good predictive validity. These results are critical in understanding the movements of cryptos and help better understand the risk associated with trading these currencies.

1. INTRODUCTION

The definition of "currency" stands for a simple and obvious meaning of coin and paper money, and it is designated as legal tender among the citizens of a country and accepted as a medium of exchange to buy and sell products or services. This definition was stated in the laws and regulations of the contusions of all countries, including the United States of America. For instance, the Code of Federal Regulations 2017 (CFR § 1010.100) supports this definition of currency. However, in recent years, a new form of trading currency has surfaced and caused a significant disturbance to established economic purposes and laws known as cryptocurrency. Starting from the idea that a non-country bound currency can be used worldwide and valued due to its limited abundance, the first cryptocurrency, Bitcoin, began in 2009 and made the news in 2012 when its value increased significantly. From this, many new cryptocurrencies have evolved and become accepted by many countries as a new form of payment. For example, in 2018, The German Ministry of Finance accepted crypto-currencies as equal to conventional means of payment. As a result, the estimated market capitalization of cryptocurrencies was approximately $200 billion at its peak in December 2019.

Since its inception, however, there has been marked volatility in these currencies compared to others, which has increased markedly since COVID-19 was declared a pandemic in 2020. For example, the exchange rate for Bitcoin peaked on December 17, 2017, when it reached an all-

Bibhu Dash et al: NLPAI, IOTBC -2024
pp. 87-106, 2024. IJCI – 2024 DOI:10.5121/ijci.2024.130207

time high of approximately $64,863.10, and has since fallen to $16,625.08 by first day of January 2023.

Other cryptocurrencies have shown similar reductions in value, and several high-profile discussions of their volatility may have further impacted their perception of the market. Previous research has focused on this potential volatility but has yet to be updated since the pandemic. For example, Chan et al., 2017 found that Bitcoin, one of the most popular and well-known currencies, had a higher average value than other cryptocurrencies but also demonstrated marked variance over their study period from 2014 to 2017. Additionally, Sapuric and Kokkinaki (2014) investigated the period preceding this and found high annualized volatility from July 2010 to August 2014 compared to other major global currencies. This may be due to Bitcoin's properties and perception as a currency, an asset, or an investment, which has been borne out in research by Kristoufek (2015) that demonstrated Bitcoin experienced both short- and long-term influence that showed both general appreciations in value over time, as well as short-term bubbles and busts.

While cryptocurrencies are a new form of payment, they are vastly different than classical cash in many aspects. One of the most important differences is their reliance on mathematical processes and statistics through mining platforms and calculated risk blockchain trading, which leads to high volatility in cryptocurrency values. Specifically, numerous mathematical equations and methods are involved in crypto statistical properties, including the algorithm for generating cryptocurrencies that may use classical hash functions and elliptic curve digital signatures algorithm. These mathematical applications are necessary to define the blockchain currencies themselves, accurately determine their value in the market, and increase confidence in their trading and use. For instance, calculating exchange rates is critical for understanding financial needs and returns. Recently, this was done by Chan et al. 2017 after examining several candidate distributions for population cryptocurrencies at the time of publication. If we understand the underlying distribution, we can have real-world impacts in calculating exchange rates, describing volatility, etc. Thus, investigating these statistical properties using fitting models is critical for a better understanding of cryptocurrencies and aiding future research in this area.

We first describe the distributions using descriptive statistics and examine the exchanges overtime during our study period to address this knowledge gap. Secondly, we fitted the top 8 cryptocurrencies using numerous distribution models, including Student t, Laplace, Skew Student t, generalized t, Asymmetric Student’s t, Inverse Gaussian, and Generalized Hyperbolic Distribution. We compared their fittings to the U.S. Dollar and Euro. Here, we show that the cryptocurrencies’ returns are non-normally distributed, and none of the distribution fitting models explains all the top 8. Third, we compared these fit statistics and reported the best-fitting models and associated parameters for these cryptocurrencies and their returns from 2017 to 2022. Finally, we generated predictive estimates for the daily log returns from January 1st to February 1st, 2023, to assess our model's accuracy on new data it was not initially trained on.

2. Aims And Objectives

we aim to study the effects of recent international events between 2017 and 2022, such as the COVID pandemic and political instabilities on the top 8 cryptocurrencies. We also provide a follow-up to previously described statistical and fitting models of these coins. In 2017, a previous report named "A Statistical Analysis of Cryptocurrencies" described the spread of the daily exchange rates and log-returns of the exchange rates of the cryptocurrencies from June 2014 until the end of February 2017. In this manuscript, we aim to provide an update on the current status of the cryptocurrency exchange and test multiple regression models for the best fit and as a follow-up for the previously published report (ref: S Chan et al. Journal of Risk and Financial

Management 2017). We also aim to demonstrate the predictive validity of these models by testing them to predict the daily log returns during the first month of 2023.

Statistics	Bitcoin	Ethereum	Tether	BNB	Solana	XRP	Cardano	DOGE	EUR*
Minimum	777.757	84.308	0.967	1.510	0.515	0.140	0.024	0.001	0.995
Q1	6,307.840	204.940	1.000	13.732	2.587	0.277	0.059	0.003	1.022
Median	9,326.688	450.407	1.001	22.495	33.584	0.387	0.139	0.003	1.054
Mean	17,564.717	1,114.652	1.002	137.479	54.869	0.532	0.498	0.060	1.055
Q3	29271.194	1834.29	1.003	294.088	91.313	0.713	0.841	0.071	1.082
Maximum	67566.828	4812.087	1.078	675.684	258.934	3.378	2.968	0.685	1.114
Skewness	1.182	1.199	2.156	1.155	1.244	2.252	1.562	2.115	-0.0523
Kurtosis	0.049	0.190	26.166	-0.127	0.493	8.614	1.669	4.907	-1.144
SD	17224.621	1226.709	0.006	185.498	64.098	0.375	0.645	0.101	0.033
Variance	296687563.07	1504815.709	0.00004	34409.613	4108.593	0.141	0.416	0.010	0.001
CV	0.981	1.1	0.006	1.349	1.168	0.705	1.293	1.703	0.029
Range	66789.071	4727.779	0.111	292.578	258.419	3.238	2.944	0.684	0.119
IQR	22963.35	1629.354	0.003	280.355	88.726	0.435	0.782	0.068	0.051

*EUR data is only from 03-03-2022 to 08-29-2022

3. DATA

The data used in this paper were obtained from Yahoo between January 1, 2017, and December 31st, 2022, using R-4.1 and the get Symbols () functions from the library (quant mod). To identify the top eight cryptocurrencies, we downloaded the latest paginated list of all active cryptocurrencies with the newest market data using the get market cap ticker all () function from the library (coin market caper). Then we sorted the cryptocurrency listing by market cap and selected the top eight, which included Bitcoin, Ethereum, Tether, BNB, Solana, XRP, Cardano, and DogeCoin. We used the same process to obtain data for January 1st to February 1st, 2023. Now we will summarize the definitions and the history of each one of the top 8 cryptocurrencies

Summary statistics of daily exchange rates of the top eight cryptocurrencies: Bitcoin, Ethereum, Tether, BNB, Solana, XRP, Cardano, and Euro versus the U.S. Dollar from 01/01/2017 to 12/31/2023

Based on our analysis, we see that Bitcoin has the most significant minimum exchange rate, at $777.76 per one Bitcoin. It also demonstrated the maximum exchange rate at $67566.828 and the highest mean exchange rate overall at $17,564.717. Tether, known to be tied to the U.S. Dollar, showed a relatively tight association with the U.S. Dollar, as expected based on its conception as a cryptocurrency. It demonstrated a minimum exchange rate of$0.967 and a maximum of $1.078, with a meaningful exchange of $1.002. This was also shown by its lowest range of all the currencies, $0.111.

Although Bitcoin and Ethereum are both the highest-valued and most conventionally popular cryptocurrencies at the time of writing (2022), we see that both have very high variability, with variances on the order of 100 million and 1 million, respectively. Not all cryptocurrencies, however, have high variability. Most notably, Tether, XRP, and Cardano have low variance. We see that not only do Bitcoin and Ethereum have high conflict and the highest ranges. Additionally, DogeCoin had one of the weakest variances of 0.010, which comports with prior work by Chan et al., from 2014 to 2017, that it similarly had low variability compared to other currencies.

We also notice that all eight cryptocurrencies are positively skewed, each with skewness greater than zero. All skewness values are relatively similar, ranging from 1.155 to 2.252. This indicates that the distribution of the exchange rate of each cryptocurrency studied is slightly skewed to the right, i.e., positively skewed. Tether has the second highest skewness value with 2.156, while XRP had the most significant skew with 2.252.

All eight cryptocurrencies also have kurtosis values that are not reasonably close to three, which is the kurtosis of a normal distribution. This means that these cryptocurrencies have different peaks than a normal distribution. Unlike Chan et al.'s (2017) investigation of the most popular cryptocurrencies from 2014-2017, where all kurtosis values are either close to 3 or greater than 3, here we see that not only are there no cryptocurrencies with kurtosis near 3, but that most cryptocurrencies have a kurtosis that is less than 3. However, it's important to note that XMP and Tether have kurtosis more significant than three. Additionally, Tether has a kurtosis of 26.17, which suggests a great deal of skew and deviation from the normal distribution. This indicates that another distribution is necessary to model it.

Since each cryptocurrency's exchange rate distribution is positively skewed, if we want to compare their average exchange rates, we should use the median to measure the center as it is less biased towards outliers. On average, we expect the exchange rates to be highest for Bitcoin, Ethereum, Solana, and BNB. Tether, XMP, and Cardano have meager exchange rates.

The differences between these results and those from the paper outlining these statistics for the most popular cryptocurrencies between 2014 and 2017 are the cryptocurrencies involved, the variability of the most popular cryptocurrencies, and the average exchange rates (as measured by the median). First, notice that the only cryptocurrency that was one of the eight most popular in 2014-2017 and 2017-2022 is Bitcoin. All the other cryptocurrencies analyzed are different, indicating that the popularity of cryptocurrencies is particularly volatile. The average exchange rates since the last paper was written have increased. In the 2014-2017 analysis, Bitcoin had the highest median exchange rate of $415.20. All other cryptocurrencies analyzed then had low median exchange rates, ranging from $0 to $4. Ethereum, BNB, and Solana each have median exchange rates over $20. This indicates that cryptocurrencies have become more popular.

Summary statistics of daily log returns of the exchange rates of the top eightcryptocurrencies: Bitcoin, Ethereum, Tether, BNB, Solana, XRP, Cardano, and Euro versus the U.S. Dollar from 01/01/2017 to 12/31/2023

Statistics	Bitcoin	Ethereum	Tether	BNB	Solana	XRP	Cardano	DOGE	EUR
Minimum	-0.465	-0.551	-0.053	-0.543	-0.465	-0.550	-0.504	-0.515	-0.011
Q1	-0.016	-0.022	-0.001	-0.023	-0.039	-0.024	-0.031	-0.025	-0.002
Median	0.002	0.001	-0.001	0.001	0.001	-0.001	0.001	-0.001	-0.001
Mean	0.001	0.001	-0.001	0.003	0.004	0.001	0.002	0.002	-0.001
Q3	0.020	0.028	0.001	0.029	0.047	0.021	0.030	0.020	0.002
Maximum	0.225	0.235	0.057	0.529	0.387	0.607	0.862	1.516	0.008
Skewness	-0.710	-0.911	0.689	0.383	-0.081	0.835	1.923	4.816	-0.542
Kurtosis	10.199	9.521	42.535	14.295	3.416	16.301	23.473	84.741	0.671
SD	0.042	0.052	0.005	0.0600	0.080	0.065	0.069	0.079	0.004
Variance	0.002	0.003	0.001	0.004	0.006	0.004	0.005	0.006	0.001
CV	28.789	58.616	-1008.222	21.278	19.961	278.492	45.704	36.937	-6.184
Range	0.69	0.786	0.11	1.072	0.852	1.157	1.366	2.031	0.019
IQR	22963.35	1629.354	0.003	280.355	88.726	0.435	0.782	0.068	0.0364

4. CORRELATION AND VARIANCE OF CRYPTOCURRENCIES

We calculated the log-returns of the exchange rates of the top eight cryptocurrencies based on market cap. In these visualizations, the distribution of log returns is approximately normally distributed for each cryptocurrency, with the average log return rate as measured by the mean for each cryptocurrency being close to zero.

5. DISTRIBUTION OF DAILY LOG RETURNS

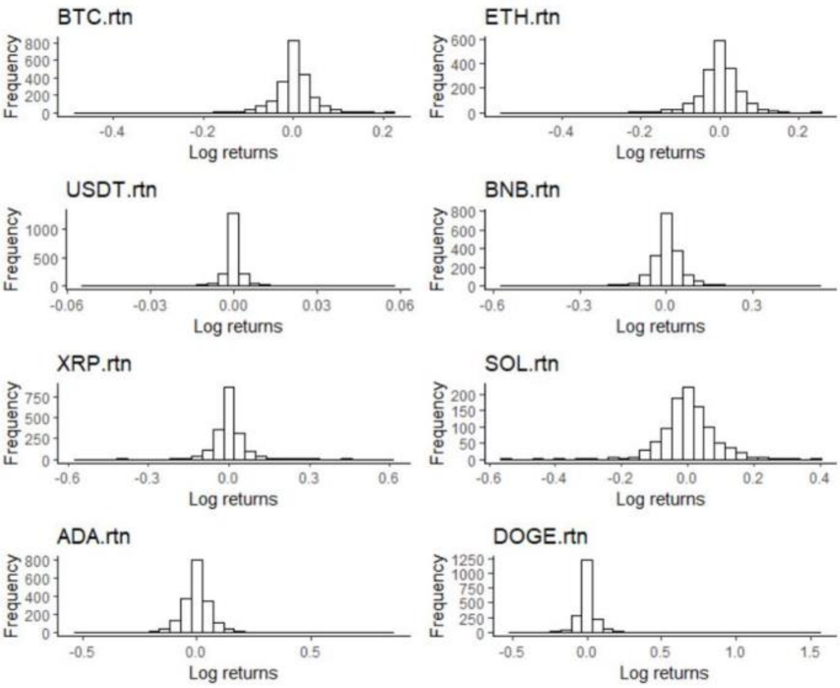

Figure 1. Histograms of daily log returns of the exchange rates of the top eight cryptocurrencies

Bitcoin, Ripple, and DogeCoin remain centered around 0 log returns, complying with prior findings from before 2017 (Chan et al.). Ethereum, Tether, Binance, Solana, and Cardano all showed that they were also centered around 0. Solana showed the most fantastic range in the distribution of log returns, with more giant tails to -0.2 and 0.2 log returns, compared to other currencies, which showed distributions more tightly centered around zero.

6. DAILY LOG RETURN FOR THE TOP 8 CRYPTOCURRENCIES

To perform distribution analyses for the Top 8 Cryptocurrencies, we first calculated the daily log return of each one and plotted them within time series (Figure 2)

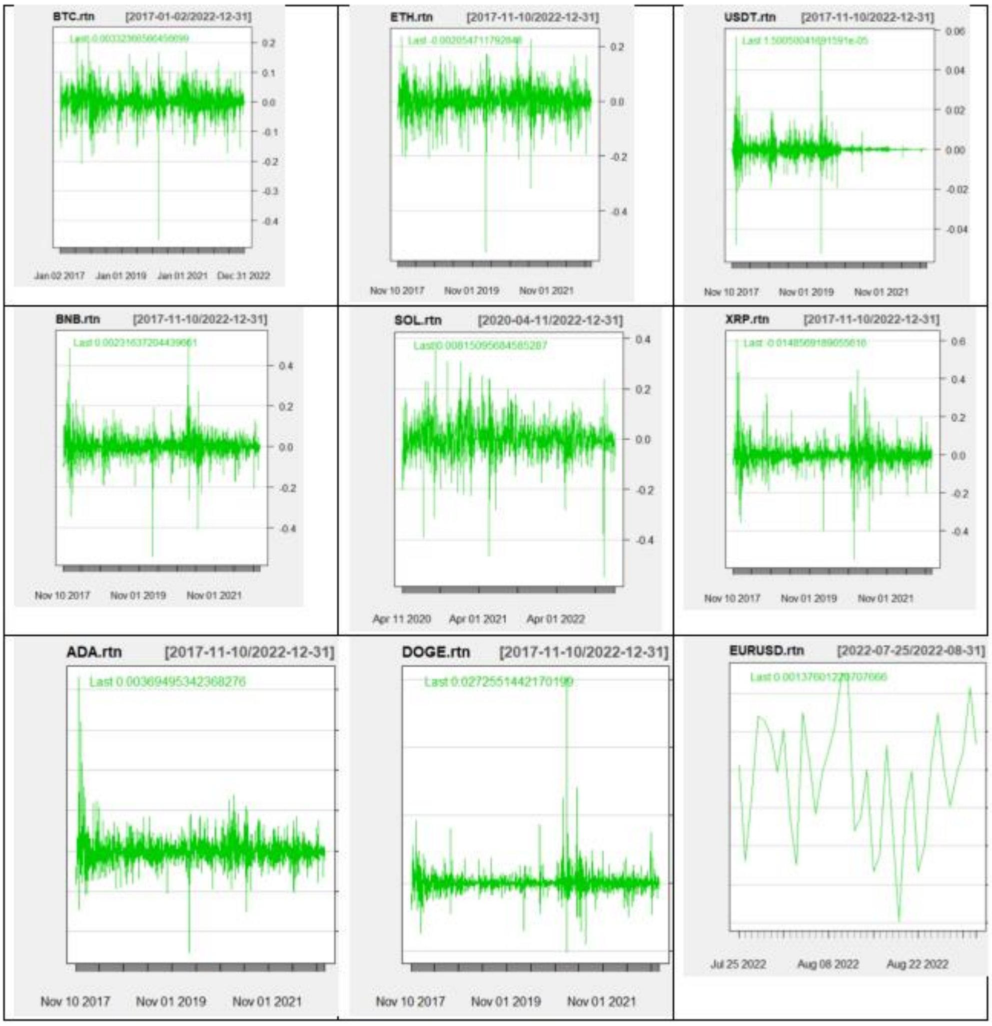

Figure 2. Daily Log Return for the Top 8 Cryptocurrencies from 01/01/2017 to 08/30/2023 and the daily log return for the EUR/USD

Figure 2 shows that the daily log returns for the top eight cryptocurrencies from 2017-2022 are centered around zero for each, though they offer some noise. Comparing this to Figure 1, we see that this visualization supports our argument of the log returns being centered about zero with some variability. Specifically, we notice that most of these cryptocurrencies had a steep dip around March 2020, at the start of the COVID-19 pandemic. At the beginning of 2020, a clear downward spike in the returns of Bitcoin and Ethereum resulted in their lowest-ever performances throughout this study. This may be related to volatility in the market due to the COVID-19 pandemic during this time.

We also notice that specific cryptocurrencies have more variability than others. Namely, Solana is highly variable. However, Tether is shown to have low variability from roughly November 2020 onwards, potentially signaling the effects of the COVID-19 pandemic on the market, which specifically for Tether is tied to the U.S. Dollar. Doge Coin also experienced relatively low variability, though there were some periods of spiking variability at the beginning of 2020. Again, the histograms in Figure 1 support these conclusions, as the histogram for Solana's log returns is more comprehensive than that of ADA. This aligns with prior estimates of Solana's variability, showing noticeable variation over time and a significant overall spread. Correlation Analysis Between the Daily Log Return for the Top 8 Cryptocurrencies

To investigate the effect of cryptocurrencies on one another, we calculated the Pearson correlation coefficient between the daily log returns of the top eightcryptocurrencies and plotted the correlationmatrix (Figure 3). The three most significant correlations between the daily log return are between the following cryptocurrencies:Bitcoin is positively correlated with Ethereum (Correlation coefficient= 0.79, p<0.001); Ethereum is positively correlated with Terra (Correlation coefficient= 0.67, p<0.001); Bitcoin is positively correlated with that of BNB (Correlation coefficient= 0.67, p<0.001).

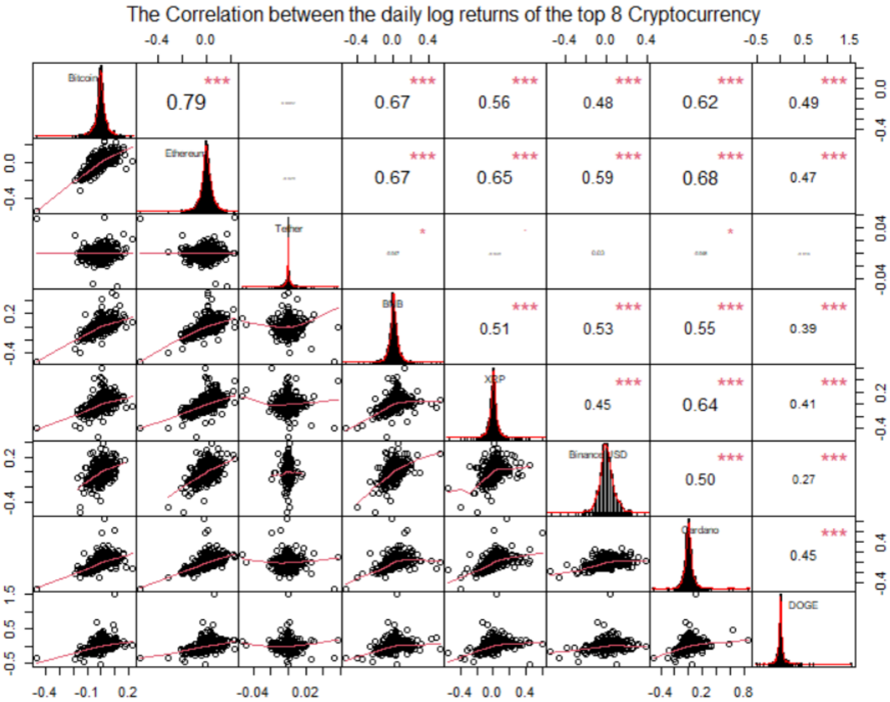

Figure 3. Visualization of the Correlation Matrix between the daily log returns of the top eight cryptocurrencies. On the top of the diagonal: the value of Person correlation plus the significance level as stars. The bivariate scatter plots with a fitted line are displayed on the bottom of the diagonal. Each significance level is associated with a symbol: p-values (0.001, 0.01, 0.05) <=>symbols ("***," "**," "*")

A central finding from this plot is that there are no negative correlations between the daily log returns of any two cryptocurrencies based on our correlation analysis. Notice that the daily log returns of Bitcoin have a strong, positive, linear relationship with most other cryptocurrencies, including Ethereum, BNB, Terra, and XRP. However, Tether's log daily returns do not correlate with any other cryptocurrency's log daily returns, potentially related to its Tether to the USD.

7. FITTING DISTRIBUTIONS AND RESULTS

To identify the best fitting distributions for the daily log returns of the top eight cryptocurrencies between the dates of January 1, 2017, and April 7, 2022, we performed the following distributions: Student t, Laplace, Skew Student t, generalized t, Zero-adjusted Inverse Gaussian, and Generalized Hyperbolic Distribution. Statistically, these distributions are defined as follows:

Let X denote a continuous random variable representing the log-returns of the cryptocurrency's exchange rate of interest. Let f (x) denote the probability density function (pdf) of X. Let F(x) mean the cumulative distribution function (CDF) of X. We suppose X follows one of seven possible distributions representing the most popular parametric distributions used in finance. They are specified as follows:

- the Student's t distribution (Cosset 1908) with

$$f(x) = \frac{K(v)}{\sigma}\left[1 + \frac{(x-\mu)^2}{\sigma^2 v}\right]^{-(1+v)/2}$$

Next, we compared the performance of the resulting fitted distributions using the goodness of fit tests, including negative log-likelihood, Akaike Information Criterion corrected for small sample sizes (AICc), and Bayesian Information Criterion (BIC) (Table 6). To fit the generalized t, skewed Student's t, and the zero-adjusted inverse Gaussian models, the package "gamlss" was used. The package "ghyp" was used to fit the generalized hyperbolic model. To check the LaPlace model, the box "VGAM" was used.

We plot the scores of the negative log-likelihood for each fitted distribution in a heatmap (Figure 4), both by column (e.g., comparing the best log-likelihood of the distributions compared to one another) and by row (e.g., comparing the best log-likelihood of the cryptocurrencies to one another). We can see that the performance of the different fitting distributions was different in various cryptocurrencies, especially Solana, which had the highest log likelihood for all fitted distributions. On the other hand, Tether had the lowest log-likelihood for 6 out of the seven fitted distributions. We also plotted the AICc and BIC scores for each fitted distribution in a heatmap (Figure 4-6). We can see that AICc and BIC are consistent with our findings in the -log likelihood.

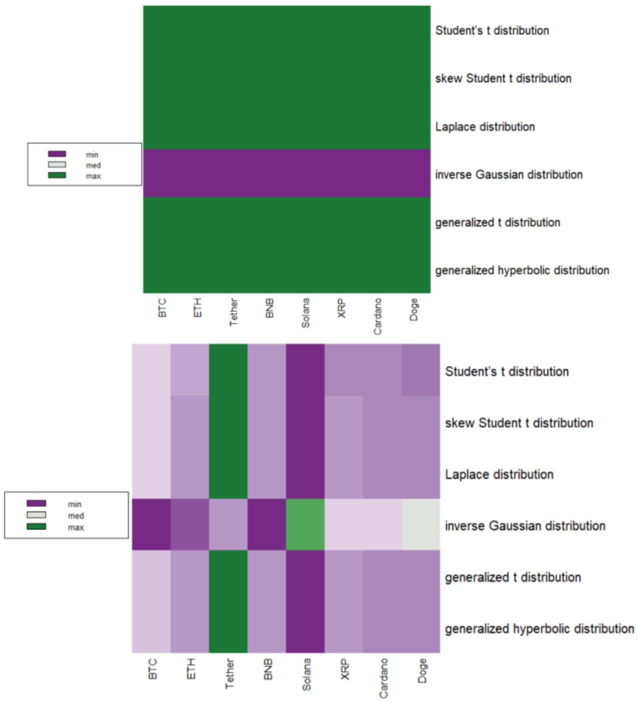

Figure 4. Heatmaps of the -log likelihood of the fitted distribution of the top 7 cryptocurrencies by column (top panel) and row (bottom panel).

The inverse Gaussian distribution across all models in Panel A shows a poor overall fit across all cryptocurrencies. The fit is so poor that it obscures the scale to understand nuances between the other models. When comparing each cryptocurrency, Tether shows a slightly better fit for the inverse Gaussian model; however, this fit remains poor compared to other models' fit with Tether. Solana performs the best across all models, indicating that it may be easiest to model compared to the others, though it shows again that inverse Gaussian demonstrates the worst fit for it.

We see that the negative log-likelihood is lowest for Tether and highest for Solana, no matter the distribution fitted to each cryptocurrency. Surprisingly, Bitcoin has the median negative log-likelihood for all distributions except for the Inverse Gaussian. Ethereum, BNB, XRP, and Cardano each have higher than average negative log likelihoods, no matter the distribution fitted. It's also interesting to note that the Inverse Gaussian yields the highest variability regarding the negative log-likelihood, as it is the maximum when provided to Bitcoin, Ethereum, and Solana but relatively low for the other cryptocurrencies.

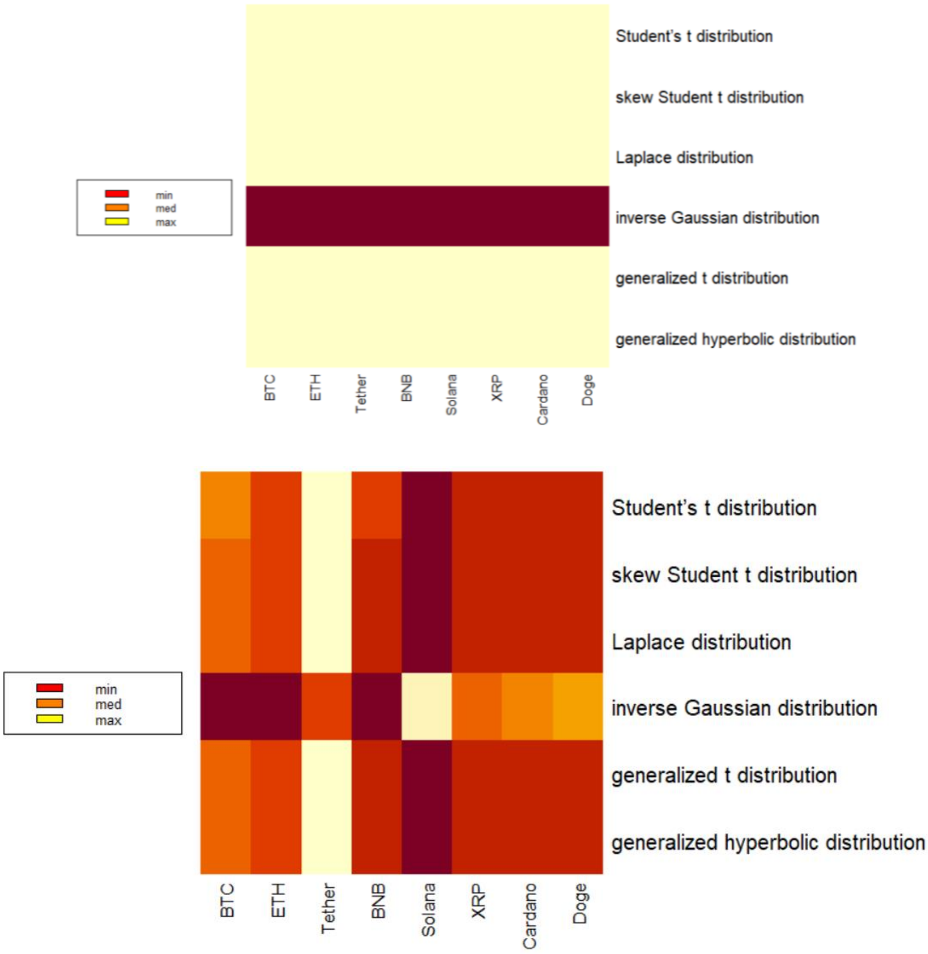

Figure 5. Heatmap of AICc of the fitted distribution of the top 8 cryptocurrencies compared by column (top panel) and row (bottom panel).

Notably, the inverse Gaussian distribution has the worst fit across all models (Panel A, which reflects comparison across columns for the best fit for each model). The performance is so poor that it obscures differences between the other models in the heatmap scale. When comparing by row (Panel B) for the best fit by each distribution, we find that most distributions performed best for Solana (indicating that Solana generally follows models compared to the other cryptocurrencies), followed by a similar performance for Binance, Ripple, Cardano, and Doge.

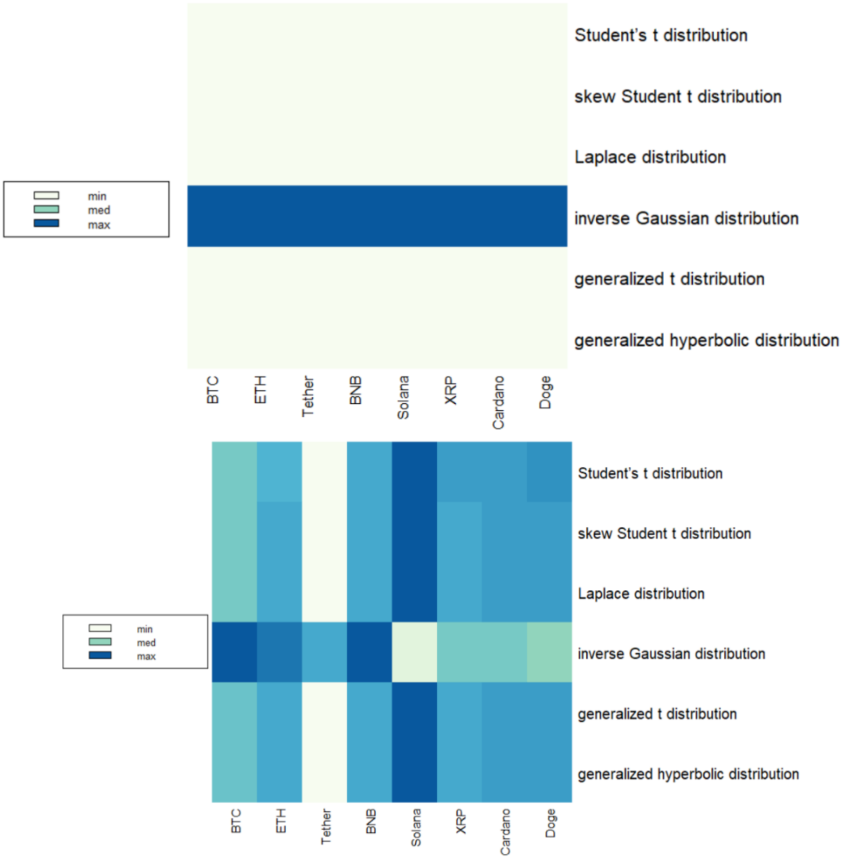

Figure 6. Heatmap of BIC of the fitted distribution of the top 8 cryptocurrencies compared by column (top panel) and row (bottom panel).

When comparing cryptocurrencies across models, we see the Inverse Gaussian needs to better fit across all. When comparing models across cryptocurrencies, Solana continues to show low BICs for models overall.

Compared to Chan et al., 2017, the best-fitting models for Bitcoin and DogeCoin remain the same. However, for Ripple (XRP), from 2017 to 2022, the Generalized t distribution exhibited the best fit compared to the normal inverse Gaussian distribution from Chan et al.'s analyses.

The best-fitting models were determined by comparing the log-likelihood formula, the AIC, and the BIC estimates for each model and seeking to maximize the log-likelihood and minimize the AIC and BIC. Based on these reasonable estimates, we report the distribution parameters (and standard errors when applicable) for the best-fitting models in our analyses.

Table 5. Best fitting distributions and parameter estimates, with standard errors given in brackets.

Cryptocurrency	**Best Fitting Distribution**	**Parameter Estimates and Standard Errors**
Bitcoin	Generalized Hyperbolic Distribution	$\hat{\mu} = 0.002$ $\hat{\gamma} =$ -0.0008 $\hat{\alpha} = 0.185$ $\hat{\sigma} = 0.041$ $\hat{\lambda} = 0.504$
Ethereum	Generalized T	$\hat{\mu} = 0.001$ (0.0009) $\hat{\sigma} = -3.183$ (0.061) $\hat{\nu} = 1.567$ (0.422) $\hat{\tau} = 0.274$ (0.104)
Tether	Generalized hyperbolic	$\hat{\mu} = 5.44$ e-7 $\hat{\gamma} =$ -5.60e-6 $\hat{\alpha} = 0.0059$ $\hat{\sigma} = 0.00404$ $\hat{\lambda} = 0.165$
BNB	Skewed T	$\hat{\mu} = 0.0007$ (0.001) $\hat{\sigma} = -3.439$ (0.033) $\hat{\nu} = 0.024$ (0.024) $\hat{\tau} = 0.933$ (0.073)
Solana	Skewed T	$\hat{\mu} = -0.003$ (0.004) $\hat{\sigma} = -2.907$(0.042) $\hat{\nu} = 0.065$ (0.056) $\hat{\tau} = 1.821$ (0.127)
XRP	Generalized T	$\hat{\mu} = -0.0008$ (0.0008) $\hat{\sigma} = -3.281$ (0.053) $\hat{\nu} = 0.717$ (0.240) $\hat{\tau} = 0.329$ (0.101)
Cardano	Generalized hyperbolic	$\hat{\mu} = -0.004$ $\hat{\gamma} = 0.005$ $\hat{\alpha} = 0.317$ $\hat{\sigma} = 0.065$ $\hat{\lambda} = -1.086$
Doge Coin	Generalized T	$\hat{\mu} = -0.001$ (0.0007) $\hat{\sigma} = -3.416$ (0.075) $\hat{\nu} = 0.890$ (0.266) $\hat{\tau} = 0.138$ (0.105)

Table 6. Fitted distributions and results for daily log returns of Bitcoin exchange rates from 01/01/2017 to 12/31/2023.

Distribution	**LL**	**AICc**	**BIC**
Student's t distribution	3888.369	-7772.74	-7761.35
Skew Student t distribution	4152.218	-8296.44	-8273.67
Laplace distribution	4168.536	-8333.07	-8321.69
generalized t distribution	4178.296	-8348.59	-8325.83
inverse Gaussian distribution	3888.369	-7772.74	-7761.35
generalized hyperbolic distribution	4177.974	-8345.95	-8340.56

Table 7. Fitted distributions and results for daily log returns of the exchange rates of Ethereum from 01/01/2017 to 12/31/2023.

Distribution	**- ln *L***	**AICC**	**BIC**
Student's t distribution	2910.492	-5816.98	-5805.91
Skew Student t distribution	3094.863	-6181.73	-6159.58
Laplace distribution	3096.509	-6189.02	-6177.94
generalized t distribution	3101.522	-6195.04	-6172.89
inverse Gaussian distribution	-829368	1658742	1658759
generalized hyperbolic distribution	3100.844	-6191.69	-6186.61

Table 8. Fitted distributions and results for daily log returns of the exchange rates of Tether from 01/01/2017 to 12/31/2023

Distribution	**- ln *L***	**AICC**	**BIC**
Student's t distribution	7492.423	-14980.8	-14969.8
Skew Student t distribution	8622.874	-17237.7	-17215.6
Laplace distribution	8373.976	-16744	-16732.9
generalized t distribution	8820.153	-17632.3	-17610.2
inverse Gaussian distribution	-669455	1338917	1338933
generalized hyperbolic distribution	8833.391	-17656.8	-17651.7

Table 9. Fitted distributions and results for daily log returns of the exchange rates of BNB from 01/01/2017 to 12/31/2023.

Distribution	**- ln *L***	**AICC**	**BIC**
Student's t distribution	2659.835	-5315.67	-5304.59
Skew Student t distribution	3003.911	-5999.82	-5977.67
Laplace distribution	2986.094	-5968.19	-5957.11
generalized t distribution	3009.552	-6011.1	-5988.95
inverse Gaussian distribution	-880549	1761104	1761121
generalized hyperbolic distribution	3005.848	-6001.7	-5996.62

Table 10. Fitted distributions and results for daily log returns of the exchange rates of Solana from 01/01/2017 to 12/31/2023

Distribution	**- ln *L***	**AICC**	**BIC**
Student's t distribution	1111.985	-2219.97	-2210.17
Skew Student t distribution	1191.594	-2375.19	-2355.58
Laplace distribution	1183.389	-2362.78	-2352.97
generalized t distribution	1191.245	-2374.49	-2354.88
inverse Gaussian distribution	-117462	234929.6	234944.3
generalized hyperbolic distribution	1192.173	-2374.35	-2370.54

Table 11. Fitted distributions and results for daily log returns of the exchange rates of XRP from 01/01/2017 to 12/31/2023.

Distribution	- ln *L*	AICC	BIC
Student's t distribution	2508.926	-5013.85	-5002.78
Skew Student t distribution	2997.638	-5987.28	-5965.12
Laplace distribution	2948.643	-5893.29	-5882.21
generalized t distribution	3003.307	-5998.61	-5976.46
inverse Gaussian distribution	-543965	1087937	1087953
generalized hyperbolic distribution	3002.669	-5995.34	-5990.26

Table 12. Fitted distributions and results for daily log returns of the exchange rates of Cardano from 01/01/2017 to 12/31/2023.

Distribution	- ln *L*	AICC	BIC
Student's t distribution	2403.578	-4803.16	-4792.08
Skew Student t distribution	2729.337	-5450.67	-5428.52
Laplace distribution	2709.453	-5414.91	-5403.83
generalized t distribution	2730.859	-5453.72	-5431.57
inverse Gaussian distribution	-514610	1029227	1029243
generalized hyperbolic distribution	2729.676	-5449.35	-5444.28

Table 13. Fitted distributions and results for daily log-returns of the exchange rates of DogeCoin from 01/01/2017 to 12/31/2023.

Distribution	- ln *L*	AICC	BIC
Student's t distribution	2117.646	-4231.29	-4220.22
Skew Student t distribution	2883.553	-5759.11	-5736.95
Laplace distribution	2787.456	-5570.91	-5559.84
generalized t distribution	2892.382	-5776.76	-5754.61
inverse Gaussian distribution	-449792	899590.2	899606.8
generalized hyperbolic distribution	2891.468	-5772.94	-5767.86

8. Quantile-Quantile Plots

We then compared each cryptocurrency's quantile-quantile (Q.Q.) plots and the best-fitting distributions. Generally, plotting the sample quantiles derived from the actual log returns against the theoretical quantiles derived from the hypothesized log returns based on our distributions, we expect a straight line of fit should there be a perfect fit of our distribution to our data.

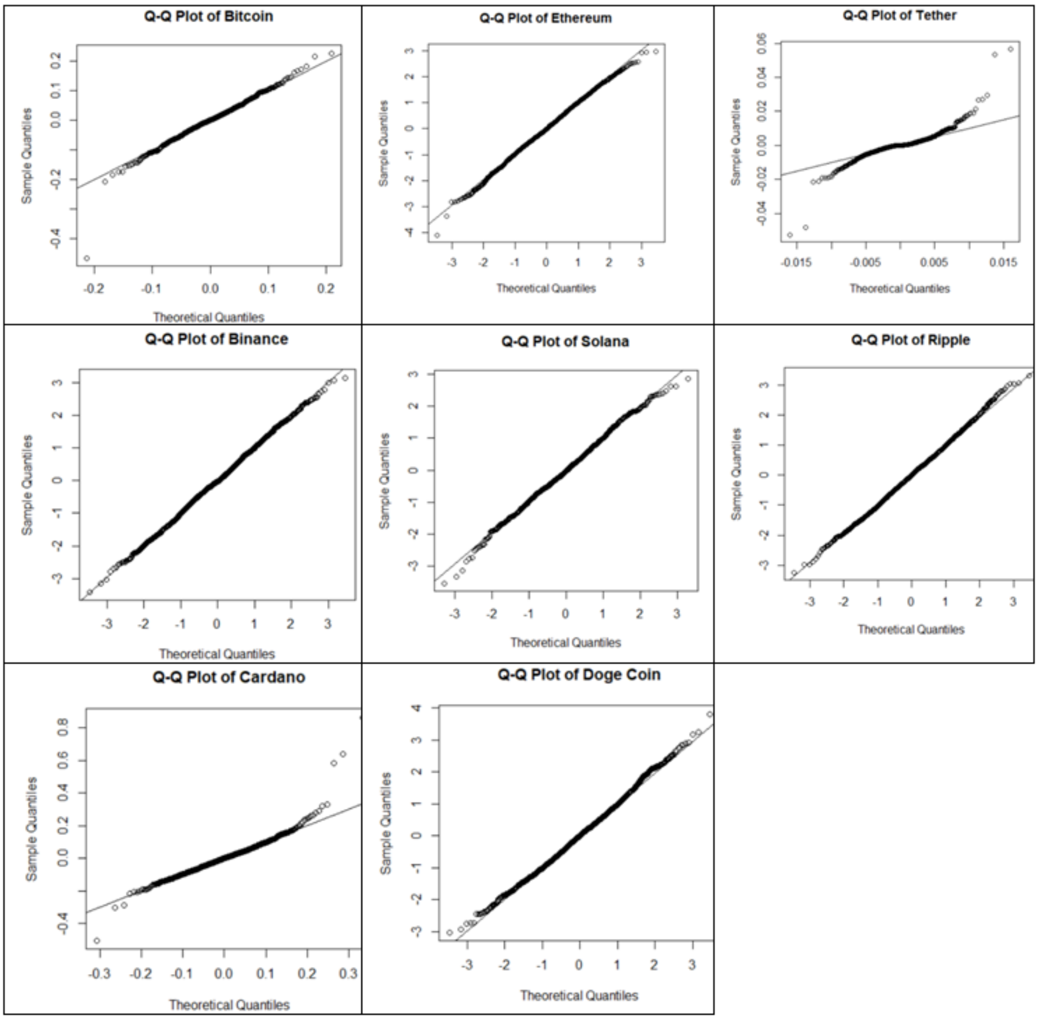

Figure 7. Q-Q Plots of the fitted distribution of the top 8 cryptocurrencies

The Q-Q plots for each of the fitted distributions of the top seven cryptocurrencies show that each is likely non-normal because although the points do the line relatively well in the middle, the projections fall off the tails on both ends. This means the sample and theoretical quantiles are not equivalent as they would be in a normal distribution, and any normality assumptions we make may be invalid.

The Q.Q. The plot for Bitcoin demonstrates both the lack of fit by the Gaussian distribution and the improved fit by the generalized hyperbolic distribution. The tails for the Gaussian distribution remain quite exaggerated from the line of the perfect fit compared to the hyperbolic distribution, which remains relatively close to the bar. While the Q.Q. The plot for Bitcoin in Chan et al., up to 2017, demonstrates poor fit in the upper tail, the Q.Q. Action for 2017 to 2022 demonstrates a good fit between expected and observed values around the perfect fit line—the Q.Q. The plot for DogeCoin up to 2017 showed poor fit and inflation in the estimates, particularly in both tails. From 2017 to 2022, we see a good fit in these tails fitting the generalized t distribution. For Ripple (XRP), up to 2017, the best-fitting model was an inverse Gaussian, demonstrating deflation in the seats in Chan et al. However, from 2017 to 2022, the best-fitting model was the generalized t-distribution, which showed slight inflation in the tails. Notably, the inflation in the seats using the generalized t distribution was less than the deflation in the bottoms using normal inverse Gaussian, suggesting a better overall fit regardless of the direction of over/under-estimation. Among the models not included in Chan et al. describing their behavior before 2017, we see slight inflation in the tails when fitting against the generalized hyperbolic distribution for Tether. For Cardano, there is similar inflation for a small number of outliers, but in the middle and most of the tails, there remains an excellent fit to the perfect fit line. BNB saw a solid fit for the skewed Student's t distribution, with little inflation or deflation. With Solana, we see deflation in the tails at higher and lower values but a good fit in the middle portion of the plot.

9. Probability- Probability Plots

We then plotted each cryptocurrency's probability-probability (P.P.) plot using our data hypothesized cumulative distribution and the cumulative sample distribution.

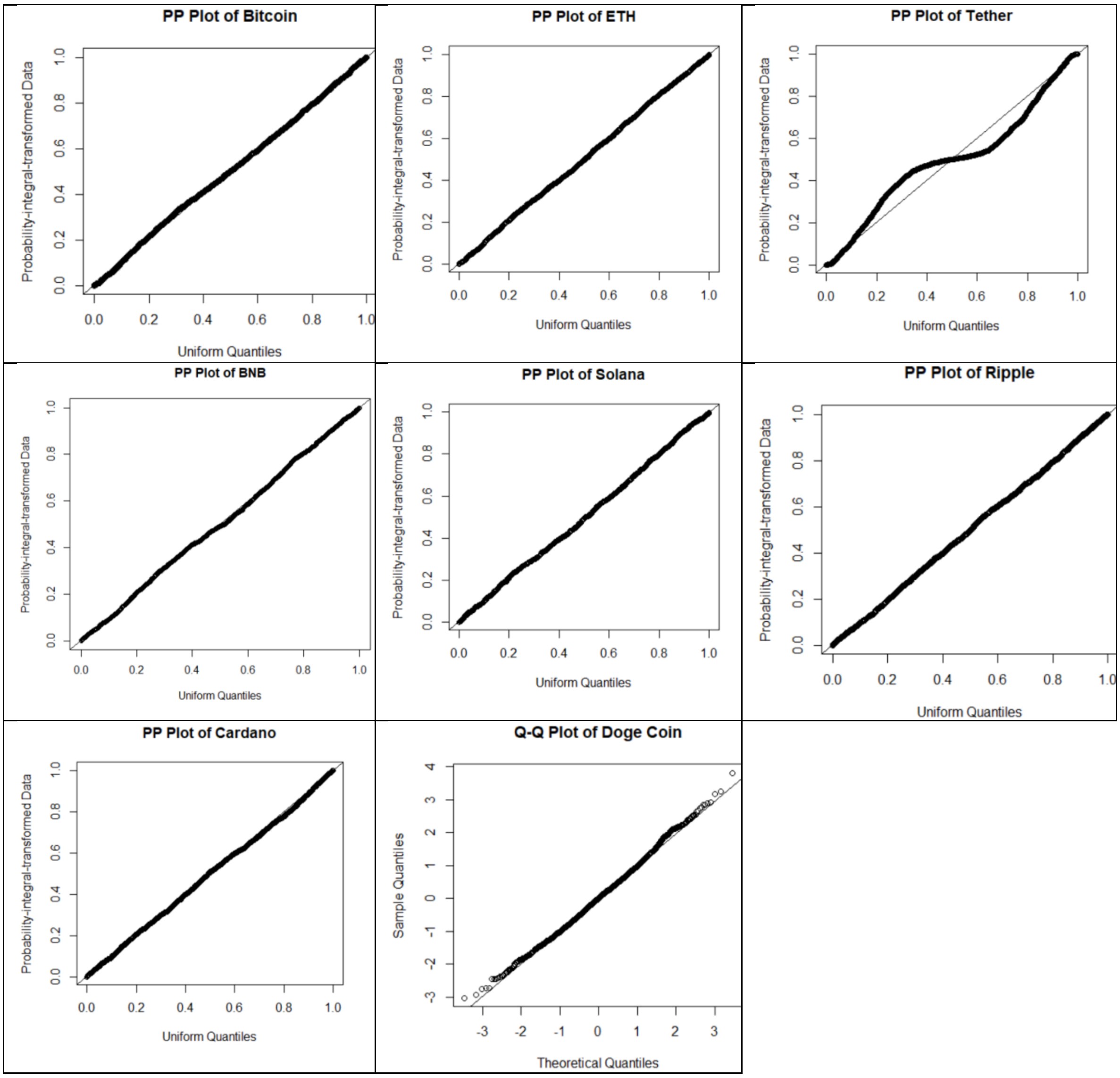

Figure 8. P-P Plots of the fitted distribution of the top 8 cryptocurrencies

The P-P plot demonstrates how closely two datasets agree by plotting their cumulative distribution functions against each other. If they agree perfectly, the points should lie in a straight line. In this case, the points generally match the straight line for all cryptocurrencies except Tether, indicating that the cumulative distributions are similar for most currencies. It is important to note that the point does pass through the line at around (0.5 0.5) for each cryptocurrency, indicating that the median of the distributions match.

Assessing Volatility and Prediction

Additionally, many volatility estimates can be conducted to describe further the change in variance of these distributions over time. We measure emotional volatility based on the rolling standard deviations (S.D.s), which calculates the standard deviation, limiting our data only to a given time window. This demonstrates whether the S.D.s are centered around a single number, indicating that there are often similar S.D.s regardless of a given time window, or if there are tails in any direction, that may mean increasing levels of volatility. For example, a uniform distribution would suggest a considerable difference in the S.D.s over a given window. In contrast, a high-peaked normal distribution would indicate that the SDs are centered around a single value.

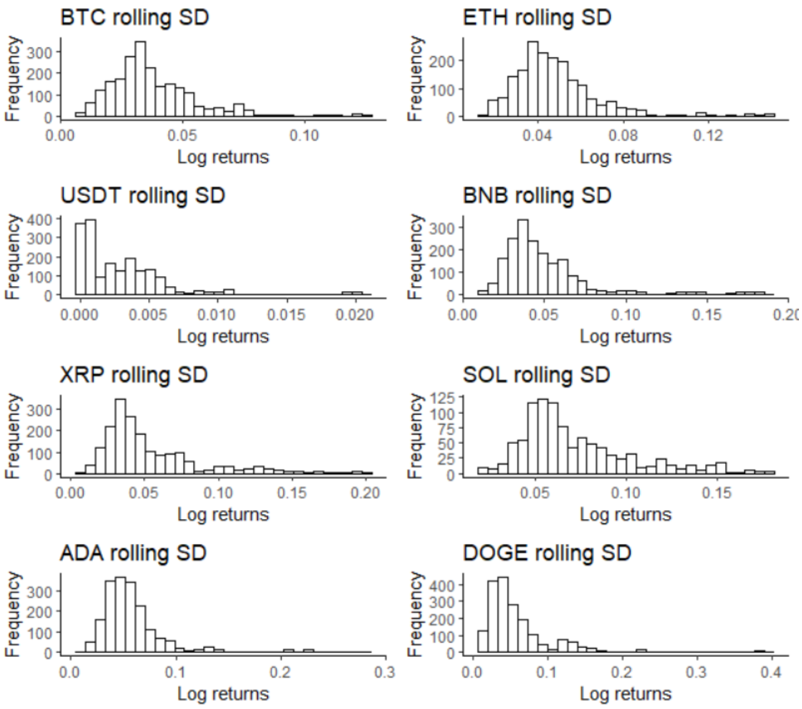

Figure 9. Rolling standard deviation distributions of the cryptocurrencies using a 20-day window.

Determining volatility via rolling standard deviations

Across the cryptocurrencies, we see that some coins demonstrate less dynamic volatility over the 20-day window due to their right-skewed distribution with high tails. However, there is a difference in the scale of this variation; Tether, for instance, has a maximum of around 0.02, while BNB's maximum rolling S.D. is near 0.2. Dogecoin has the highest full SD near 0.40, indicating that, for a handful of 20-day windows, there was a large amount of volatility compared to its usual S.D**.**

Applying Models for Prediction

Finally, we sought to apply these models to predict the daily log returns for the first month of 2023 and examine their accuracy using data from January 1st to February 1st. Understanding this predictive validity is important to estimate the future returns of these currencies and as a potential proof of concept for evaluating the returns of cryptocurrencies more broadly.

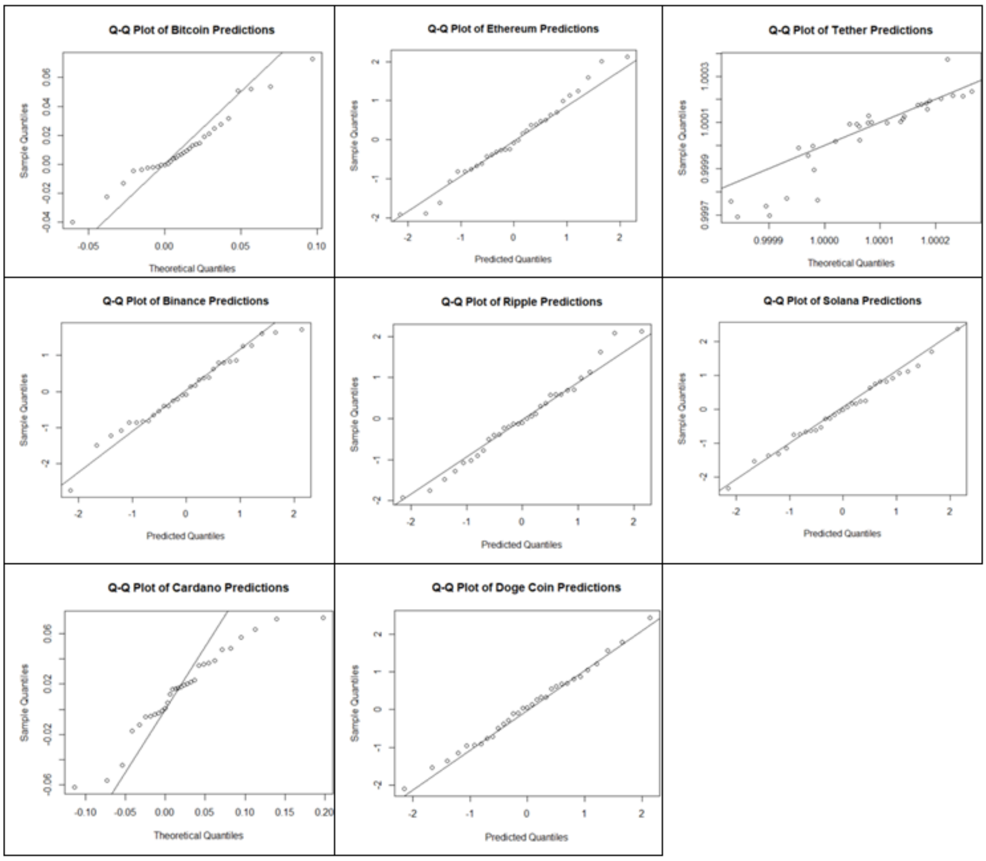

Figure 10. Q-Q plots of Predicted vs. Sample Values for January 1, 2023, through February 1, 2023, for the Top 8 Cryptocurrencies.

There is slight evidence of inflation in our predicted values for January 1st to February 1st, 2023, using our models trained on data from 2017 to 2022 for Ripple and Ethereum, as evidenced by the left and right tails being slightly below and slightly above the line of perfect fit, respectively. However, we find evidence of good fit for Doge Coin, Solana, and Binance using our models trained in the previous 5 years data (2017 to 2022). This suggests we can use our trained models to predict further daily log returns for future dates with the parameters in Table 5. In our models that used the generalized hyperbolic distribution, there was generally a worse fit with the evidence of inflation in Bitcoin, Tether, and Cardano predictions for the first month of 2023. Tether, in particular, had the worst performance, with little clear line throughout the QQ plot. Cardano showed some evidence of a line. However, there was a noticeable overestimation in the higher quantiles. Bitcoin similarly remained closer to the perfect fit line, though this also had some evidence of inflation in the first tail. Given these three currencies had worse performance than the generalized T and skewed T models, this may suggest that, for smaller time windows, such as one month compared to the past five years, the generalized hyperbolic has poor performance but, on average, over longer time windows, performs better.

10. CONCLUSION

Using eight of the most popular parametric financial distributions, we have analyzed the exchange rate of the top seven cryptocurrencies versus the Euro and U.S. Dollar. Our analysis of over five years of data shows that most cryptocurrencies exhibit heavy tails. Using the discrimination criteria of the log-likelihood, AICc, and BIC, we found that no single distribution tested yields the best fit jointly across the data for all seven cryptocurrencies. This is consistent with the previous paper examining this problem with data from 2014-2017.

Before 2017, Chan et al. found the generalized hyperbolic distribution best fit Bitcoin, which this analysis also found from 2017 to 2022, and the generalized t distribution best fit Doge Coin, which this analysis also found from 2017 to 2022. The generalized hyperbolic distribution is defined by the standard variance-mean mixture with the generalized inverse Gaussian distribution, which means that while the normal inverse Gaussian distribution is a particular case of this generalized family of distributions, only the generalized form fits the Bitcoin data well.

Notably, before 2017, Chan et al. found the normal inverse Gaussian distribution best fit Ripple, but this analysis found the generalized t distribution demonstrated a better fit. The normal inverse Gaussian distribution is a subset of the hyperbolic distribution, which is suited for particularly heavy tails, similar to the generalized t distribution, which also can model heavy-tailed data. Despite this, the two distributions are not especially related (e.g., the normal inverse Gaussian is not a special case of the generalized t distribution, or vice versa).

Notably, the zero-adjusted inverse Gaussian distribution performed the worst across all models. First, the inverse Gaussian distribution, on its own, cannot model zero returns, which is why zero adjustments was required. Additionally, it must be better suited for modeling negative values, which can occur with daily log returns. It may have been better fitting in Chan et al. due to having few or no negative log returns over their study period because of the rising value of these cryptocurrencies. During the COVID-19 pandemic, however, there was increased volatility and reduced returns compared to prior, which may have led to the inverse Gaussian distribution no longer being a suitable distribution candidate for these cryptocurrencies due to violation of its assumptions (particularly a non-negative dependent variable).

These results are significant in risk management, where it is essential to know the risks of certain investments, how they are correlated with others, and how their values may change over time. Our study generally found good predictions for most cryptocurrencies, though the generalized hyperbolic distribution seemed to have a poorer fit in this small time window.

This study mirrors a similar investigation examining the top seven cryptocurrencies from 2014 to 2017, which yielded identical results.

There are many possible extensions of this work. Namely, we are interested in using multivariate models to examine joint distributions of log returns and investigating nonparametric models for cryptocurrency log returns.

REFERENCES

[1] Chan, S., Chu, J., Nadarajah, S., & Osterrieder, J. (2017). A statistical analysis of Cryptocurrencies. Journal of Risk and Financial Management, 10(2), 12. https://doi.org/10.3390/jrfm10020012

[2] Sapuric, S., & Kokkinaki, A. (2014). Bitcoin is volatile! Isn't that right? Lecture Notes in Business Information Processing, 183, 255–265. https://doi.org/10.1007/978-3-319-11460-6_22

[3] Kristoufek, L. (2015). What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. PLOS ONE, 10(4), e0123923. https://doi.org/10.1371/journal.pone.0123923

[4] Bitcoin Project 2022. Frequently Asked Questions. Available online: https://bitcoin.org/en/faq#what-is-bitcoin (accessed on October 6, 2022).

[5] Ethereum 2022. What is Ethereum? Available online: https://ethereum.org/en/what-is-ethereum/ (accessed on October 6, 2022).

[6] Tether 2022. What is Tether? Available online: https://www.investopedia.com/terms/t/tetherusdt.asp (accessed on October 6, 2022).

[7] BNB 2022. What is BNB? Available online: https://www.investopedia.com/terms/b/binancecoin-bnb.asp (accessed on October 6, 2022).

[8] Solana 2022. What is Solana? Available online: https://www.investopedia.com/solana-5210472 (accessed on October 6, 2022)

[9] Ripple 2022. What is Ripple? Available online: https://ripple.com/faq/ (accessed on October 6, 2022)

[10] Cardano 2022. What is Cardano? Available online: https://www.investopedia.com/cardanodefinition-4683961 (accessed on October 6, 2022)

[11] Doge Coin 2022. What is DogeCoin? Available online: https://coinmarketcap.com/currencies/dogecoin/

[12] Fidelity 2022. What is Market Cap? Available online: https://www.fidelity.com/learningcenter/trading-investing/fundamental-analysis/understanding-market-capitalization (accessed on October 6, 2022)