Can Bitcoin become a viable alternative to fiat currencies?

Econometrics Project

Course: Insurance and Econometrics - 8 CFU

Maggioni Fabio

Scaramuzza Riccardo

Veneziani Luca

A.Y. 2020/2021

POLITECNICO
MILANO 1863

Index

1 Introduction

The work is based on a paper, submitted by Vavrinec Cermak to Skidmore College in may 2017, of title "Can Bitcoin Become a Viable Alternative to Fiat Currencies? An empirical analysis of Bitcoin's volatility based on a GARCH model".
In 2017, Bitcoin already established itself as a diffused and reliable mean of payment: most of the doubts and complaints (both from the economic community and common people) aroused in the first years of living of the Bitcoin, were dismantled. Bitcoin became instead one of the hot topic in the scientific community, both from a structural and economic point of view.
The cited work by Cermak addressed an interesting question, wondering if Bitcoin could play the role of a fiat currency: at first, he considered a theoretical (economical) approach, based on the Austrian and Keynesian schools, identifying the common points and differences between the two. The conclusion was, actually, that the former could not be considered as a fiat currency for several reasons, but identify one of the most important problem in the Bitcoin (high level) volatility.
A first important thing to note is that the variable of interest is not the straight Bitcoin price (BTC), but the Bitcoin Price Index (BPI). Actually, in the early days of Bitcoin it was difficult to determine the exact value of a coin, as the value of Bitcoin versus other currencies depended itself on the exchange considered. This prompted the cryptocurrency news organization, Coindesk, to fix the situation by creating the Bitcoin Price Index (BPI). This index polls the largest, most established cryptocurrency exchanges, aggregates the data, and then applies the aggregated data to show a balanced and realistic picture of Bicoin's value.
In order to analyze such feature, Cermak relied on a classical ARMA-GARCH model, which was enhanced with an exogenous part both in the mean and in the variance equation[1], introduced due to the fact that economic literature identifies in macroeconomic variables (such as inflation, interest rates, money supply, exports and GDP, etc) the driving factors of fiat currency volatility. In order to take into account the decentralized nature of the Bitcoin, which can be considered, actually, as an international currency, a stack of macroeconomic indicators was taken per each one of the countries where it was traded the most (in order, China, US, Europe and Japan). Those stacks comprehend: the currency exchange rate (US $ are taken as a baseline), the stock market index, 10-year government bond yield, and a three month interbank rate. Moreover, also the price of gold is suggested by literature, because of the shared qualities with bitcoin, such as limited quantity, easy transportation, easily division, the inability to counterfeit, and acceptance as barter.
Our work is structured to reproduce the analysis carried out by Cermak, but develops deeper each single part of the study: first focusing on the ARMA modeling of BPI, then on the exogenous part, to finally evaluate different models to put things togheter. A remarkable attention is put on the training of ARMAX-GARCHX models, carried

[1] Hereafter, called ARMAX-GARCHX model

out through custom-implemented MLE methods.
Two further remarks must be pointed out: the considered time span is actually the one from the original study (August 18, 2010 to March 17, 2017), chosen as we try to keep things consistent with it. In spite of that, sources of our data are different due to data availability. We downloaded BPI from Coindesk in accordance with Cermak, but we retrieved Shangai Index, SP500 and Eurostock from Yahoo Finance and the rest from FRED - Federal Reserve Economic Data, while his sources were FRED and Trading Economics. This lead to some slight differences in the results.

2 BPI Analysis

We started by analyzing the Bitcoin Price Index series, which is the price of one Bitcoin in USD and uses an average from the world's leading bitcoin exchange. We decided to analyze the log price of the BPI index: in this way, two equivalent price changes are represented by the same vertical distance on the scale. First of all, the Q-Q Plot of the log prices showed us that they are not normally distributed (Figure 1). Actually, on the Q-Q Plot, data appear S-shaped meaning that their distribution is under-dispersed relative to a normal distribution (reduced numbers of outliers).

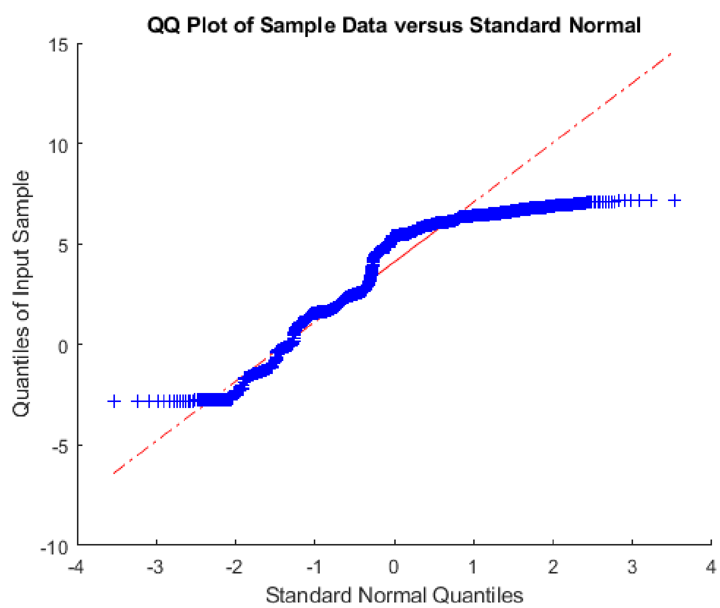

Figure 1: Q-Q plot of the BPI's log prices

We plotted in Figure 2 the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) in order to check for the stationarity of the time series . In particular, if the time series contains a unit root then it is nonstationary. From the plot it is clear that the ACF decays to zero very slowly: this was the first signal of non stationarity. Also the plot of the PACF shows that probably the log prices of the BPI are a unit root series.

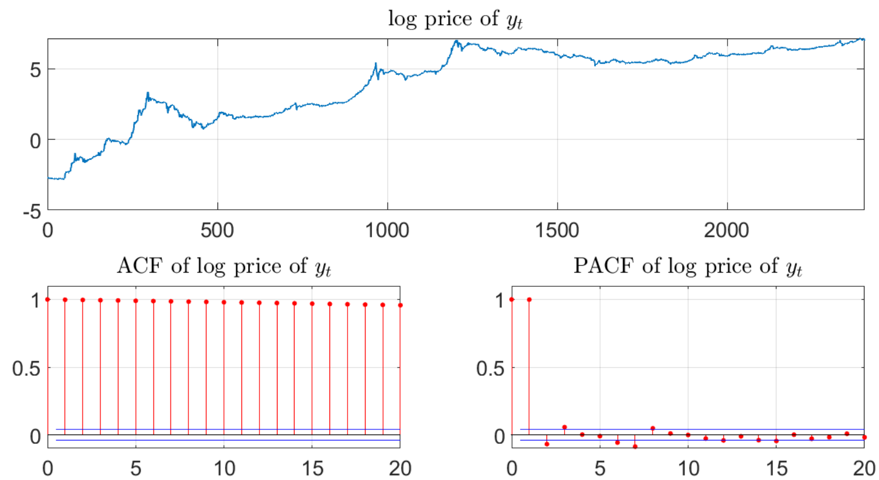

Figure 2: Log price of BPI, ACF and PACF

We used a statistical hypothesis unit root test to objectively determine whether the series requires differencing. In our analysis, we used the Augmented Dickey Fuller test. The test is done by using the built-in function of MatLab *adftest*.

1. The first test was a simple Dickey-Fuller test with 0 lag and trend stationary model (TS).

   $$Null\ Model : y(t) = c + y(t-1) + e(t)$$

   $$Alternative\ Model : y(t) = c + d \cdot t + a \cdot y(t-1) + e(t)$$

   The pValue of the test is very high (0.4877), meaning that we have no evidence to reject the null hypothesis of unit root at 1% or 5% level. Moreover, we accepted the trend equal to zero at 1% and 5% level (pValue of 0.3604).

2. The second test was an Augmented Dickey-Fuller test with 8 lags and trend stationary model (TS). In particular, for the selection of the lag length we used the general-to-specific methodology: we started with a long lag length (8 in this case). In case the t-statistics was insignificant, the next step would require to estimate the regression using the lag length decreased by 1.

   $$Null\ Model : y(t) = c + y(t-1) + \beta_1 \cdot y(t-2)..\beta_8 \cdot y(t-9) + e(t)$$

   $$Alternative\ Model : y(t) = c + d \cdot t + a \cdot y(t-1) + \beta_1 \cdot y(t-2)...\beta_8 \cdot y(t-9) + e(t)$$

The pValue of the test is very high (0.4104) and so we have no evidence to reject the null hypothesis of unit root at 1% or 5% level. Moreover, we accepted the trend equal to zero at 1% and 5% level (pValue of 0.2581). The relevant lags with 1% level were: 1, 2, 5, 6, 7.

3. The last test was an Augmented Dickey-Fuller test with 7 lags (since from the previous test the relevant lags were up to lag number 7) and using an autoregressive with drift model (ADR) since from the previous test we had evidence for the trend to be zero.

   $$Null\ Model : y(t) = y(t-1) + \beta_1 \cdot y(t-2)..\beta_7 \cdot y(t-8) + e(t)$$

   $$Alternative\ Model : y(t) = c + a \cdot y(t-1) + \beta_1 \cdot y(t-2)...\beta_7 \cdot y(t-8) + e(t)$$

   In this case the pValue is much lower than before (0.0393) so we can accept unitary root only at level 1%. Moreover, we reject the null hypothesis of drift equal zero (the pValue is 0). Also, all the lags are significant.

In conclusion, despite the pValue decreased a lot in the last test, overall we had evidence that the series of the BPI log prices has a unit root. This means that the series is not stationary. In this series, the effects of shocks do not disappear over time, eventually leading to instability. From the theory, we know that if a series with a unit root problem has d roots equal to 1 then the d^{th} order difference of the sequence will generate a stationary series. For this reason, we moved to the log returns of the BPI series.

$$\Delta y(t) = log(P_t) - log(P_{t-1}) \tag{1}$$

with: P_t = price of BTC in USD at time t

Moreover, the adoption of log returns is very common in the financial world as they posses the log-normality property. As previously done, we started with the Q-Q Plot: data appears as a flipped-S shape (Figure 3). This means that this series presents an increased number of outliers. From the plots of the Autocorrelation Function and the Partial Autocorrelation Function we do not have the suspect of unit root's presence anymore (Figure 4). Moreover, the two plots show that a possible seasonality effect is present (lags 5 and 6). To be sure that the log returns of BPI are a stationary series we performed the Augmented Dickey-Fuller test.

1. We started with a simple Dickey-Fuller test with 0 lags and trend stationary model (TS). Since the pValue is 0.001 we can reject the null hypothesis of unitary root. Moreover, we accepted at a level of 5% the trend equal to zero.

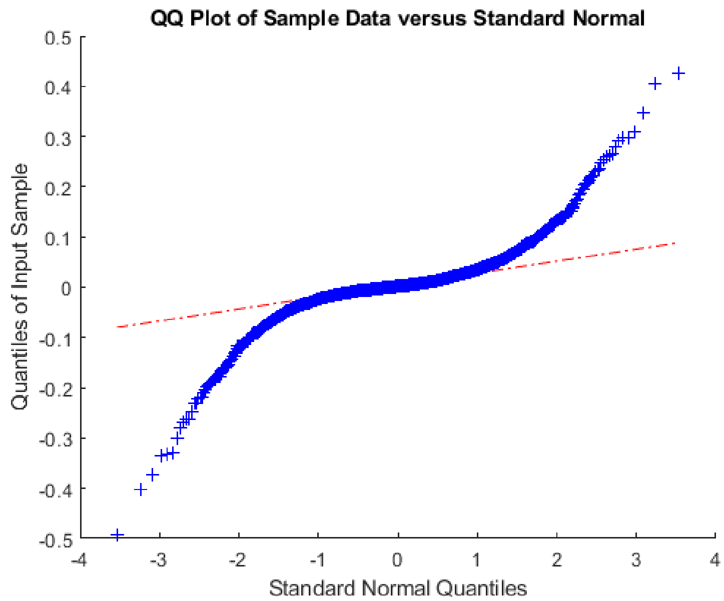

Figure 3: Q-Q plot of the BPI's log returns

2. Always using a trend stationary model (TS), we performed an Augemented Dickey-Fuller test with 7 lags. As before, we reject the null hypothesis of unitary root with over 0.1% level and we accept the trend equal to zero at 5% level. At a 5% level the relevant lags are: 2, 3, 6.

3. In the end, since from the previous test the relevant lags were up to lag number 6, we performed an Augmented Dickey-Fuller test with 6 lags. From the previous test we had evidence for the trend to be zero so we changed the model from TS to ADR (autoregressive with drift mode). The pValues are respectively 0.001 and 0.0032 so we have evidence to reject the null hypothesis of unitary root and drift equal zero.

Since all the Augmented Dickey-Fuller tests on the log returns gave us the same result we can conclude that this series is stationary.

3 ARMA

At this point, since we have a stationary time series, we are able to go deeper in the analysis. The conditional mean equation of the stationary time series must always be represented by an AR, MA or ARMA model.
In the financial world, and in the particular analysis of the volatility of financial returns, the AR model is widely used. For this reason, the writers of the paper decided to use an AR(1) model for the conditional mean equation. Nevertheless, we decided to perform a model selection procedure in order to find the best model. We estimated

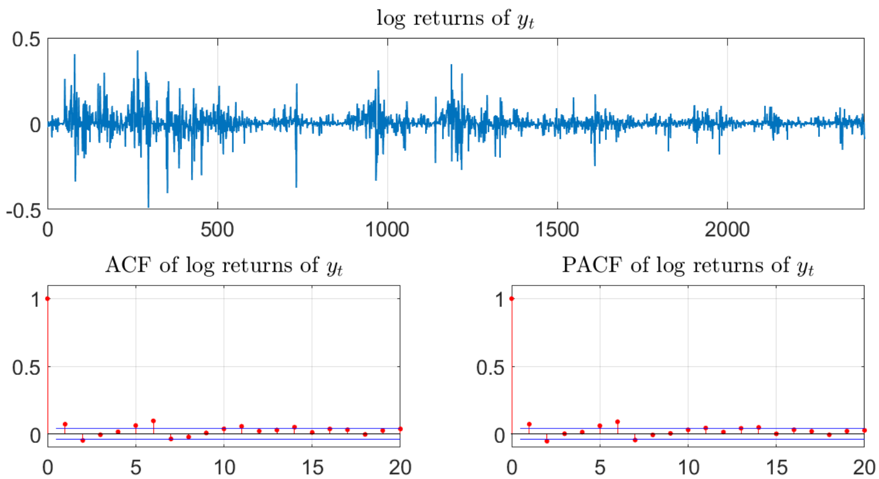

Figure 4: Log returns of BPI, ACF and PACF

ARMA models with AR and MA orders ranging from 0 to 6, considering all possible combinations. During the estimation process, we also computed the loglikelihood of each model and obtained the relative Akaike's and Bayes's Information Criterions (AIC and BIC). These indices aim at finding a parsimonious model which is able to properly fit data, minimizing a trade-off function between number of calibrated parameters and loglikelihood.
The model with the lowest BIC is the ARMA(0,1) while the model with the lowest AIC is the ARMA(0,6). Following the $Parsimony Principle$, i.e. parsimonious models produce better forecast than overparametrized ones, we selected the ARMA(0,1) model. This means that the log returns of the BPI time series are described by a Moving Average model of order 1. After the model selection procedure, we tested the MA(1) model for ARCH effects. The ARCH test assesses the null hypothesis that a series of p residuals exhibits no conditional heteroskedasticity (ARCH effects)[2].

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \cdot \epsilon_{t-1}^2 + ... + \alpha_p \cdot \epsilon_{t-p}^2 + v_t$$
$$Null\ Hypothesis : \alpha_1 = \alpha_2 = ... = \alpha_p = 0$$
$$Alternative\ Hypothesis : \alpha_k \neq 0 for\ at\ least\ one\ k = 1, ..., p$$

The test returns a pValue of zero and for this reason we rejected the Null Hypothesis of no ARCH effects. Consequently, we enhanced our MA(1) with a GARCH model, which is one of the most performing models to analyze and forecast volatility. In particular, following the suggestion of the paper, we used a GARCH(1,1) model since it is widely used in currency volatility estimation. So, we fitted the variance of our best model with a GARCH(1,1) and re-estimated the model parameters. Nevertheless, performing the ARCH test, we still obtained a pValue equal zero.
Since the authors of the paper used an AR(1) model for the mean equation, we decided

[2] We always consider p=1

to also estimate an AR(1)-GARCH(1,1) model. Still, the same result for ARCH effect was obtained in this case.

4 Linear Regression

After the modeling of the dependent variable, we addressed our attention to the exogenous ones. As already mentioned in the Introduction, these are chosen according to a specific criteria: we look for macroeconomic variables believed to affect the volatility of fiat currencies of countries where Bitcoin is traded the most. However, even if their primary role is related to volatility explanation, it is important to test their influence also from a regression point of view, as our final aim is to build a comprehensive ARMAX-GARCHX model. Of course, if Bitcoin shares the same nature and behaviour of fiat currencies, we expect that such variables would not have great explanatory power [3].

	N	Mean	SD	Min	Max
BPI	2404	271.31512	291.30473	0.059	1290.786
CNY_USD	1648	0.15713	0.00531	0.14371	0.16555
EUR_USD	1648	1.25727	0.11887	1.0375	1.4875
JPY_USD	1648	0.01043	0.00167	0.00796	0.0132
GOLD_USD	1664	1378.75147	201.58808	1050.6	1896.5
SHANGAICompInd	1716	2694.13108	594.0808	1950.01196	5166.35009
S&P500	1716	1716.96868	360.11263	1047.21997	2395.95996
EUROSTOCK	1716	2909.49044	380.86677	1995.01001	3828.78002
NIKKEI	1614	13863.9889	3930.1547	8160.01	20868.03
CHINA10Y	1727	3.59759	0.46525	2.66	4.71
USA10Y	1646	2.27949	0.50212	1.37	3.75
GERMANY10Y	1765	1.30897	0.89626	-0.184	3.497
Japan10Y	1690	0.59998	0.38816	-0.291	1.351
3MShibor	1642	4.17323	0.98490	2.4654	6.4611
US3MLibor	1664	0.40444	0.21666	0.22285	1.15178
EUR3MLibor	1678	0.32656	0.55146	-0.35714	1.56
JPY3MLibor	1664	0.12430	0.07753	-0.076	0.23875
Timestamp	2404	735569.5	694.119	734368	736771

Table 1: Summary statistics

The macroeconomic indicators identified are: the currency exchange rate with respect to US $ (representing macroeconomic strength of a country) plus the price of gold , the stock market index (representing stock market), 10-year government bond yield (representing bond market), and a three month interbank rate (representing bank sector). Therefore the complete list of exogenous variables considere are: CNY/USD,

[3]from a classical regression point of view, i.e. explaining the conditional mean

	Lags H1	P-value
CNY/USD	2	0.1282
ShangaiCompInd	6	0.5964
China10Y	5	0.9629
3MonthShibor	9	0.6126

Table 2: ADF test: China related exogenous variables

EUR/USD, JPY/USD, Gold Price per Ounce; Shanghai Composite Stock Market Index, S&P500, EuroStoxx 50, Japan NIKKEI 225 Stock Market Index; China Government Bond 10Y, United States Government Bond 10Y, Germany Government Bond 10Y (as a proxy for a 10Y bond of the European Union), Japan Government Bond 10Y; China SHIBOR Three Month Rate, US Dollar LIBOR Three Month Rate, Euro LIBOR Three Month Rate and Japanese Yen LIBOR Three Month Rate. In addition, time was considered initially as a regressor, using Matlab timestamp format.
Clearly, each of those time series showed unavailable values, corresponding to different financial markets closures, differently to the BPI series. To deal with that, we adopt a row-wise removal approach.
In Table 1, instead, are reported summary statistics for each explanatory variable. Slight discrepancies between data used in the paper and our data can be spotted.
It is necessary to highlight that we did not take into account these variables in raw format, but performed some transformations. In particular:

1. **Exchange rates** and **Price of Gold** : Logarithmic returns
2. **Stock market Indexes**: Logarithmic returns
3. **10Y Bonds**: Differences
4. **Interbank rates**: Differences
5. **Time**: Standard normalization

The main reason why we applied these transformations was to stabilize also the regressors' time series, which, as the BPI series, seems to be integrated. To ensure that, we firstly looked at ACF and PACF of each series. As an example, in Figure 5 are reported those relative to CNY/USD: just noticing the slowly decreasing behaviour of the former was enough to wonder if a unit root was present; note, moreover, that such behaviour was similar per each regressor (with the exception of Time).
Finally, an Augmented Dickey-Fuller test was carried out per each variable, where the number of lags (characterizing the alternative hypothesis) considered in each test is inferred looking at the corresponding plots. The results confirmed that we could not reject the unit root hypothesis in all cases. As an example, in Table 2 are reported the pValues of the test performed on the China-related variables.

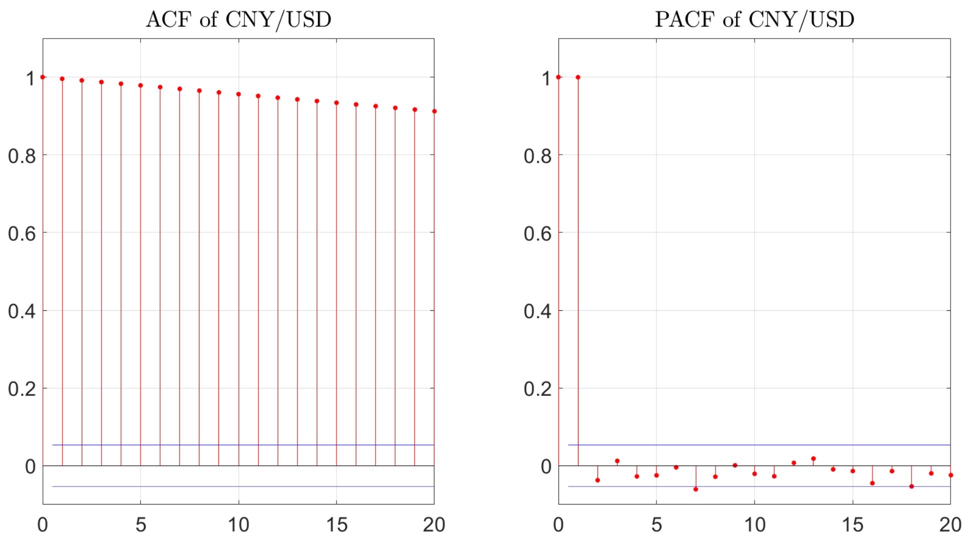

Figure 5: CNY/USD: ACF and PACF

Once assessed the integrated nature of the exogenous time series, we decided to apply different transformations according to the different nature of financial data considered. Note that the "Time" variable was standardized, instead, simply to avoid that its magnitude could cause instability in the regression; however, it was included for completeness and would be discarded in future modeling, thus its importance was relative. Another relevant issue, at this stage, was whether or not to consider the (obtained after transformations) exogenous time series lagged by one period. This strategy was proposed in the paper, due to the assumption of delayed feedback; we actually evaluated both the alternatives, and finally decided to stick with this assumption.[4]
It is important to highlight that, using lagged regressors, if we were able to produce a significant and explanatory powerful regression model, this would have lead to a chance of arbitrage; thus, what we expected to find was actually the opposite.

Firstly, we looked at simple scatterplots, using BPI as the dependent variable, to grasp some insights: as expected, the dependencies seems more relevant from a volatility point of view that from a conditional mean one: all the plots showed a strong heteroskedastic behaviour; in Figure 6 are reported the scatterplots relative to market stock indexes as example.

Next step was to check the correlation matrix. We did that with a double aim: confirm that regressors are little correlated with the dependent variable and assess possible multicollinearity. In Figure 7 is reported the heatmap of the correlation matrix: looking at the first row, the first hypothesis seemed to be true: actually, the maximum correlation in absolute value between BPI and a regressor (japan10Y) is -0.09. Also the hypothesis of multicollinearity between subgroups of regressors seemed to be plausible,

[4]For this reason, in the following are showed results relative to lagged exogenous variables

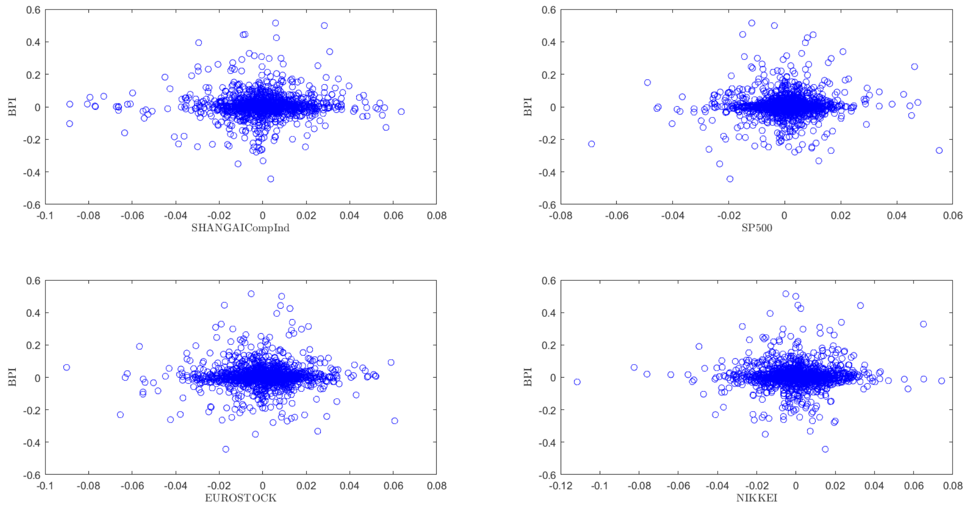

Figure 6: Scatterplots: Stock Market Indexes vs BPI

possibly leading to problems in coefficients estimation.

Finally, we estimated the linear regression model. In Table 3 are reported the coefficients estimates, Standard error estimates,T-statistics and p-values of the one-at-a-time significance tests. The achieved R2 and adjusted R2 are respectively 0.0208 and 0.00893, testifying the little explanatory power. In addition to that, the classical significance F test reported a p-value of 0.0292 and the most significant coefficient turned out to be the constant one (Intercept). Afterwards, we performed different tests on the model residuals, whose results are reported in Table 4.
Firstly we checked normality of residuals. Already from the residual QQ-plot, it appeared that the residual distribution was too fat tailed. The non normality was confirmed via a Shapiro-Wilk test. Second step was to check for residuals heteroskedasticity. Both Fitted/Residuals and Scale/Location plots suggested an heteroskedastic behaviour, which was confirmed at all significance levels through Breusch-Pagan and White tests. Finally, we focused on residuals autocorrelation performing Durbin-Watson, Ljung-Box and Breusch-Godfrey test: all gave positive results of autocorrelation, spotting the necessity of time-dependent (ARMA) modeling of BPI.
Once assessed heteroskedasticity and autocorrelation, we computed robust estimators for the coefficients standard errors,in order to validate our previous considerations on the coefficients significance. In last two columns of Table 3 are reported robust T statistics and p-values for the one-at-a time (robust) significance test.

	Estimate	SE	tStat	pValue	tStat [robust]	pVal [robust]
Intercept	0.0067	0.0020	3.3533	0.0008	2.9171	0.0036
CNY_USD	0.3208	1.3438	0.2388	0.8113	0.2291	0.8189
EUR_USD	0.4129	0.3703	1.1151	0.2650	1.0313	0.3026
JPY_USD	0.1519	0.3735	0.4069	0.6841	0.3902	0.6965
GOLD_USD	0.1515	0.1852	0.8184	0.4133	0.7614	0.4465
ShCompInd	-0.0067	0.1312	-0.0514	0.9590	-0.063	0.9498
S&P500	0.2076	0.2715	0.7646	0.4447	0.7149	0.4748
EUROSTOCK	-0.0424	0.1950	-0.2176	0.8278	-0.2146	0.8301
NIKKEI	0.2692	0.1590	1.6935	0.0906	1.6776	0.0937
CHINA10Y	-0.0071	0.0359	-0.1985	0.8428	-0.1919	0.8479
USA10Y	0.0261	0.0495	0.5275	0.5979	0.5739	0.5661
GER10Y	-0.0064	0.0562	-0.1140	0.9093	-0.1236	0.9016
JPY10Y	-0.3670	0.1049	-3.4993	0.0005	-2.7076	0.0069
3MShibor	-0.0173	0.0388	-0.4446	0.6567	-0.4387	0.6610
US3MLibor	0.1351	0.5456	0.2476	0.8044	0.3285	0.7426
EUR3MLibor	0.3959	0.2937	1.3474	0.1780	1.0611	0.2888
JPY3MLibor	-0.2108	0.5942	-0.3548	0.7228	-0.4191	0.6752
Timestamp	-0.0046	0.0020	-2.2911	0.0221	-1.8853	0.0596

Table 3: Coefficients and their significance

	pVal
Shapiro-Wilk	0
Breusch-Pagan	1.4475e-07
White	4.5992e-05
Durbin-Watson	1.0125e-04
Breusch-Godfrey	7.7707e-08
Ljung-Box	6.4524e-08

Table 4: Residuals testing

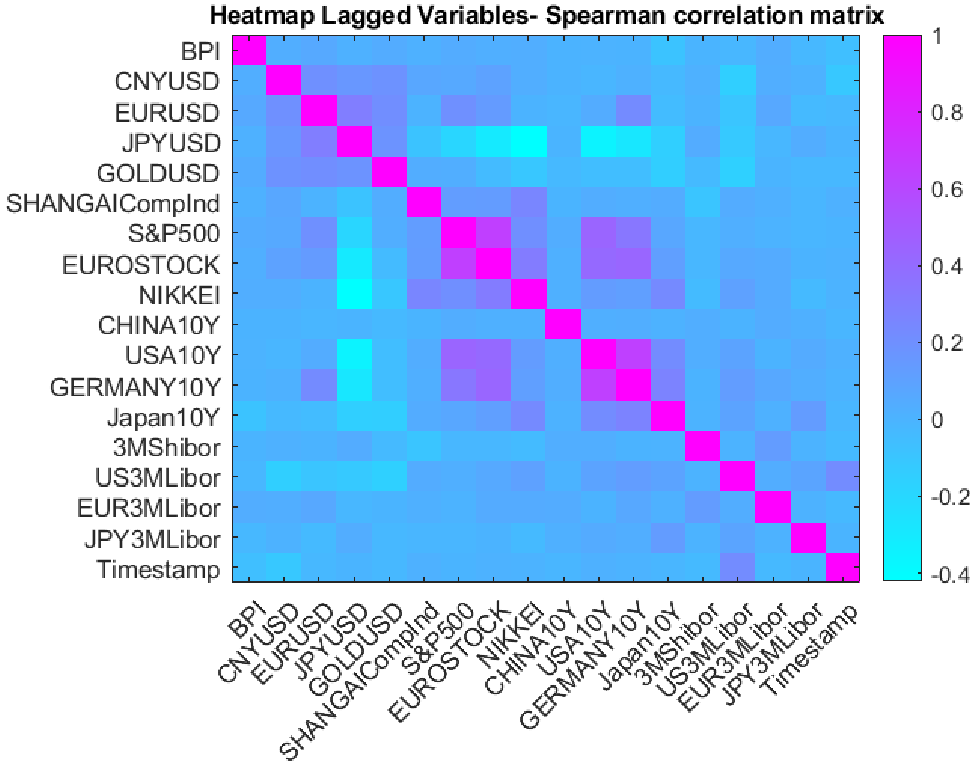

Figure 7: Heatmap of correlation matrix

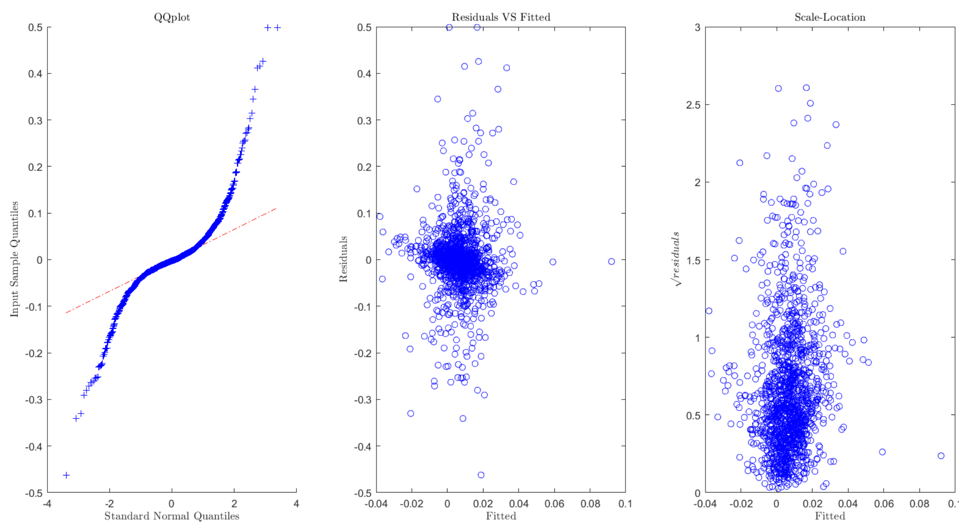

Figure 8: Residuals plots

5 Armax Garchx

Our aim is to study the possibility for Bitcoin to become a viable alternative to fiat currency. In order to achieve this goal, we have to analyze if the explanatory variables (exchange rates, stock market indexes, 10 years government's bonds, 3 months interbank rates) affect Bitcoin's volatility. We started by introducing the explanatory variables in the mean equation, implementing an ARMAX model. We conduced again a model selection based on AIC and BIC values, to find out which AR and MA orders has higher performances. It should be pointed out that, in accordance with what is done in the paper and with the linear regression study, the explanatory variables are lagged by one period. It means that, if variables are statistically significant, this could present an opportunity of arbitrage.
The model with the lowest BIC and AIC is ARMAX(3,6).
We conducted an ARCH test in order to see if there are ARCH effects. We obtained a pValue of 0 as result of this test, and for this reason we rejected the Null Hypothesis of no ARCH effects. Moreover, while looking at the plot of the Autocorrelation function and the Partial Autocorrelation function of the squared residuals, we can see significant autocorrelation in the residuals, confirming the result of the ARCH test.

At this point we tried to model the variance using two different models:

1. The first attempt was done by modeling the variance with a GARCH(1,1) model. We performed another model selection to find out the best and most parsimonious ARMAX-GARCH(1,1) model. According to the Bayes's Information Criterion, the ARMAX(4,2)-GARCH(1,1) was selected. Nevertheless, the problem of ARCH effects is not solved: the pValue of the ARCH-test is of the order of magnitude of 10^{-14}.
   In Table 5, the estimated parameters and their significance are reported. In particular, we noticed that most part of the exogenous variables are significant. This suggests that this model is not able to produce accurate estimations: in fact, according to the paper, the explanatory variables should be statistically insignificant, otherwise it could present an opportunity for arbitrage. Probably, this happens because of the complexity of the model, the presence of a lot of exogenous variables and the multicollinearity between some of them. Furthermore, from the financial point of view we can point out that, for this model, variables like the CNY/USD or the China 3 months Interbank rate are insignificant: this is unexpected since, as previously mentioned, Bitcoin is most traded in China. Moreover, by looking at the plot (restricted to observations from 500 to 1000) of the BPI's log returns and its confidence interval at 95% in Figure 9 obtained with the ARMAX(4,2)-GARCH(1,1), we can notice that we are not able to describe our time series with this model. Most of the real values are in fact outside of the confidence interval obtained. We can also point out that the variance of the model is very small and therefore the Confidence Interval is very tight.

2. The second attempt was done by modeling the variance using an EGARCH(1,1).

Parameter	Estimated value	Estimated pVal
Constant β_0	0.0034752***	0.0029777
AR(1)	0.98871***	6.739e-266
AR(2)	-1.0272***	1.7861e-117
AR(3)	0.077309*	0.073249
AR(4)	-0.010879	0.70943
MA(1)	-0.9449***	0
MA(2)	1***	0
$\Delta lnCNY_{t-1}$	-0.16235	0.27253
$\Delta lnEUR_{t-1}$	-0.15539***	6.5571e-05
$\Delta lnJPY_{t-1}$	-0.11768***	0.0032873
$\Delta lnGold_{t-1}$	-0.017936	0.41167
$\Delta lnSSI_{t-1}$	0.048509***	0.00013877
$\Delta lnS\&P500_{t-1}$	-0.014493	0.63474
$\Delta lnStoxx_{t-1}$	0.065258***	0.0029748
$\Delta lnNIKKEI_{t-1}$	-0.079461***	0.00022433
$\Delta ChinaB_{t-1}$	-0.057145***	1.4625e-15
ΔUSB_{t-1}	-0.024909***	9.7724e-09
$\Delta GerB_{t-1}$	0.026836***	6.2752e-08
$\Delta JapB_{t-1}$	-0.040501***	0.0006135
$\Delta ChinaI_{t-1}$	0.0070757	0.18691
ΔUSI_{t-1}	-0.24059***	6.4646e-05
$\Delta GerI_{t-1}$	-0.068478**	0.036369
$\Delta JapI_{t-1}$	-0.071262	0.38019

Table 5: ARMAX(4,2)-GARCH(1,1) coefficients

An interesting feature of asset prices is that "bad" news seems to have a more pronounced effect on volatility than does "good" news. For this reason, a positive shock will have a smaller effect on volatility than a negative shock with the same magnitude. The EGARCH model is able to capture this asymmetric effect of news that is not contemplated by the GARCH model.

The equation of the conditional variance is in log-linear form. This model is introduced to relax the non-negativity constraints: in fact, a problem of the GARCH is that all of the estimated coefficients need to be positive.

In this case, we obtained a model with less parameters than before(ARMAX(1,1)-EGARCH(1,1)) but the changes in the model that describe the conditional variance do not affect the result: the pValue of the ARCH-test is still zero and most of the realization fall out of the 95% CI. Despite the lower number of significant parameters with respect to the ARMAX-GARCH, the EGARCH model is still unable to solve the parameters' significance issue.

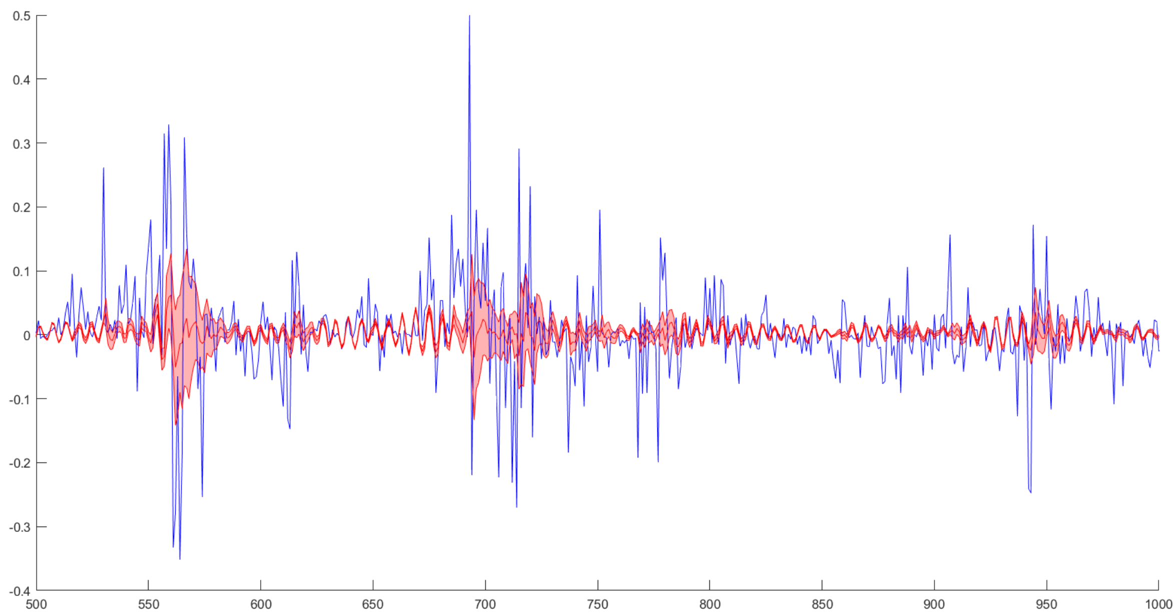

Figure 9: Log returns of BPI vs the estimated 95% CI using an ARMAX(4,2)-GARCH(1,1)

Being both the pValues of the two ARCH-tests almost zero, we concluded that the GARCH model we are using is not appropriate. A possible solution would consist in trying different variants of GARCH model (IGARCH, TGARCH, APARCH etc.) with different lag orders. However, since we wanted to be consistent with the paper, we implemented an ARMAX-GARCHX model.

The ARMAX-GARCHX model further develops the variance estimation of the ARMAX-GARCH model, inferring the variance not only on the base of previous steps variance estimates and residuals, but also accounts for the exogenous variables used in the mean equation. To ensure positivity of the variance, we modeled the exogenous contribution as the exponential of a linear combination of the exogenous variables. Moreover, it is important to point out that we pre-process the exogenous variables in a different way before feeding them to the mean equation and to the variance equation: log returns of exchange rates, gold price and stock indices were considered in both mean and variance equations, while plain 10Y government bonds and interbank rates were taken for the variance equation, and their differences for the mean one. The GARCH order chosen was still GARCH(1,1), as in the previous, again due to its consolidated use in literature. For the ARMA part, we implemented two different models: a pure autoregressive model, as done in the paper, and an ARMA(4,2), in line with the previous study of ARMAX models of different orders. The equations of the ARX(1)-GARCHX(1,1) model are as follow:

$$
\begin{aligned}
\sigma_t^2 =& exp(\lambda_0 + \lambda_1 \Delta lnCNY_{t-1} + \lambda_2 \Delta lnEUR_{t-1} + \lambda_3 \Delta lnJPY_{t-1} + \lambda_4 \Delta lnGold_{t-1} + \\
& \lambda_5 \Delta lnSSI_{t-1} + \lambda_6 \Delta lnS\&P500_{t-1} + \lambda_7 \Delta lnStoxx_{t-1} + \lambda_8 \Delta lnNIKKEI_{t-1} + \\
& \lambda_9 ChinaB_{t-1} + \lambda_{10} USDB_{t-1} + \lambda_{11} GerB_{t-1} + \lambda_{12} JapB_{t-1} + \\
& \lambda_{13} ChinaI_{t-1} + \lambda_{14} USI_{t-1} + \lambda_{15} GerI_{t-1} + \lambda_{16} JapI_{t-1} + \\
& \gamma \sigma_{t-1}^2 + \alpha \epsilon_{t-1}^2 + \epsilon_t
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
Y_t =& \beta_0 + \beta_1 \Delta lnCNY_{t-1} + \beta_2 \Delta lnEUR_{t-1} + \beta_3 \Delta lnJPY_{t-1} + \beta_4 \Delta lnGold_{t-1} + \\
& \beta_5 \Delta lnSSI_{t-1} + \beta_6 \Delta lnS\&P500_{t-1} + \beta_7 \Delta lnStoxx_{t-1} + \beta_8 \Delta lnNIKKEI_{t-1} + \\
& \beta_9 \Delta ChinaB_{t-1} + \beta_{10} \Delta USB_{t-1} + \beta_{11} \Delta GerB_{t-1} + \beta_{12} \Delta JapB_{t-1} + \\
& \beta_{13} \Delta ChinaI_{t-1} + \beta_{14} \Delta USI_{t-1} + \beta_{15} \Delta GerI_{t-1} + \beta_{16} \Delta JapI_{t-1} + \\
& \delta Y_{t-1} + \epsilon_t
\end{aligned}
\tag{3}
$$

The extension to the ARMAX(4,2)-GARCHX(1,1) is not reported, but straightforward.
In order to calibrate this model, we could not rely on built-in MATLAB functions, as no one allowed the presence of an exogenous part in the variance equation,thus we implemented it on our own. First we defined a function which computes the negative log likelihood of the model assuming Gaussian distribution. Then we implemented a minimization problem to minimize the negative log likelihood with the parameters of the model as minimization variables. In this way we obtain the MLE of the model. Due to the high number of parameters (37 and 42) the problem presented several local minima. In order to find the global minimum, or at least a particularly good local one, we performed an extensive sensitivity analysis. First of all we exploited the built-in MatLab function *arima* and *estimate* to fit an ARX(1)-GARCH(1,1) (or ARMAX(4,2)-GARCH(1,1)). This calibration is fundamental to have a meaningful initial guess of all parameters of the mean equation, and of the GARCH parameters of the variance equation. Then we proposed 13 different starting points for the initial guess of the exogenous parameters of the variance equation. We propose both guesses with the same value for each exogenous parameter and guesses where the initial value is different for each parameter, varying in a certain range. The sensitivity analysis showed that two of the 13 problems converge to the same minimum point, except for numerical error differences, associated to the lowest negative log likelihood with respect to the overall results.

The individuated calibration of the ARX(1)-GARCHX(1,1) is different from the calibration presented in the paper. However, testing the calibration proposed in the paper on our data we retrieve a way higher negative log likelihood, than the one obtained on our calibration. This is surely due to the use of slightly different data, as pointed out at the beginning of the report, combined with the difficulty of the minimization problem; nevertheless, we would expect a certain similarity in the results.

The first proof of this expected similarity is that, using the paper's calibration as initial guess of our optimization model, the algorithm easily converge to the same optimal solution found also by the sensitivity analysis. To further study the similarity between the two calibration, we compare which variables are significant in the two models, and if these variables have the same sign. In order to estimate the significance level of the estimated parameter, we performed a likelihood ratio test (LRT). The rationale of this test is to evaluate the impact of restricting one or more parameter to 0 on the model likelihood. Specifically the test can be computed as:

$$LRT = -2ln\frac{sup_{\theta\in\Theta_0}\mathcal{L}(\theta)}{sup_{\theta\in\Theta}\mathcal{L}(\theta)} = 2[sup_{\theta\in\Theta_0}l(\theta) - sup_{\theta\in\Theta}l(\theta)] \tag{4}$$

with $\mathcal{L}$ likelihood and l negative log likelihood
This statistic test is distributed as a χ^2 with as many degrees of freedom as the number of restricted parameters. In order to obtain a good estimate of this test, we needed an accurate estimate of the restricted negative log likelihood. For this reason, we solved several times each restricted optimization problem with different initial guesses. To increase the probability of converging to the restricted global minimum, we insert among the initial guesses, the optimal solution of the unrestricted problem, the optimal initial guess found in the first sensitivity analysis and the calibration proposed by the paper. We decided in this case to propose 5 different initial guesses to obtain a meaningful trade-off between accuracy and computational feasibility.[5]
In table 6 we showed the estimated parameters compared with the ones reported in the paper and their significances.

Observing the table we recognized several common traits among the two calibrations and we could also analyze their economic interpretability. First of all we could observe that most of the mean equation parameters have a low significance in our calibration. This is consistent with the paper results and with the "no arbitrage" principle. Moreover this result supported our skepticism with respect to the high significance estimated for the ARMAX-GARCH model by MATLAB. With respect to the paper results, our most relevant regressor was the 10Y Bond issued by Chinese Government. This can be related to the fact that Bitcoin was mainly traded against CNY and so we expect higher dependence from Chinese macroeconomic variables.
Considering the variance, we notice that it was significantly related with several regressors. This happens both in Cermak model and in our one. Moreover, when the same coefficient was relevant in both models, it also has the same sign, except for the US Interbank rate. Always in accordance with the paper we saw that Japanese macroeconomic variables have a lower impact, and this can be again justified with the lower volume of Bitcoin trades in Japan, with respect to the other main economies, over the considered period.
Plotting the realization of the time series against the estimated series, we obtained Fig-

[5] In this way performing the LRT estimation for all parameters, one by one, requires almost 20 minutes

Parameter	Paper value	Estimated value	Estimated pVal
$Constant\ \beta_0$	0.0309**	0.0025**	0.0140
$\Delta lnCNY_{t-1}$	-0.6110	-0.1847	0.8023
$\Delta lnEUR_{t-1}$	-0.7133***	-0.1507	0.4719
$\Delta lnJPY_{t-1}$	0.2073	-0.0340	0.8639
$\Delta lnGold_{t-1}$	0.1738	0.1908*	0.0873
$\Delta lnSSI_{t-1}$	0.0820	-0.1485**	0.0457
$\Delta lnS\&P500_{t-1}$	0.0802	-0.0606	0.6762
$\Delta lnStoxx_{t-1}$	0.1361	0.2587**	0.0147
$\Delta lnNIKKEI_{t-1}$	-0.1218	-0.0842	0.2971
$\Delta ChinaB_{t-1}$	-0.0044	-0.0610***	0.0084
ΔUSB_{t-1}	-0.0030	0.0080	0.7659
$\Delta GerB_{t-1}$	0.0106	0.0426	0.1562
$\Delta JapB_{t-1}$	-0.0090	-0.0742	0.1710
$\Delta ChinaI_{t-1}$	-0.0015	-0.0092	0.7268
ΔUSI_{t-1}	-0.0094	-0.2117	0.3939
$\Delta GerI_{t-1}$	-0.0011	-0.2510	0.2868
$\Delta JapI_{t-1}$	-0.0302	0.2543	0.2493
$AR\ \delta$	0.0196	0.0569*	0.0871
$Constant\ \lambda_0$	-6.7890***	-21.4034***	0.0000
$\Delta lnCNY_{t-1}$	-199.9000***	-196.7888***	0.0035
$\Delta lnEUR_{t-1}$	61.1000***	196.5913***	0.0008
$\Delta lnJPY_{t-1}$	3.6040	13.9993	0.6694
$\Delta lnGold_{t-1}$	-21.2400***	-34.2596**	0.0156
$\Delta lnSSI_{t-1}$	39.9100***	11.9778	0.3437
$\Delta lnS\&P500_{t-1}$	36.0800***	-4.8813	0.8068
$\Delta lnStoxx_{t-1}$	-38.9500***	-16.0684	0.2915
$\Delta lnNIKKEI_{t-1}$	-7.7370	12.4091	0.4051
$ChinaB_{t-1}$	1.1590***	4.6542***	0.0000
USB_{t-1}	-2.1010***	-2.1035	0.3086
$GerB_{t-1}$	4.5500***	2.8707***	0.0001
$JapB_{t-1}$	-9.6600***	-5.7913***	0.0015
$ChinaI_{t-1}$	-0.2688***	-0.8106**	0.0159
USI_{t-1}	-1.7260*	5.1058***	0.0000
$GerI_{t-1}$	2.0420***	1.6551***	0.0011
$JapI_{t-1}$	3.6060	3.0724	0.7001
$GARCH\ \gamma$	0.3660***	0.7671 ***	0
$ARCH\ \alpha$	0.1924***	0.2027***	0

Table 6: Mean parameters above, variance parameters below

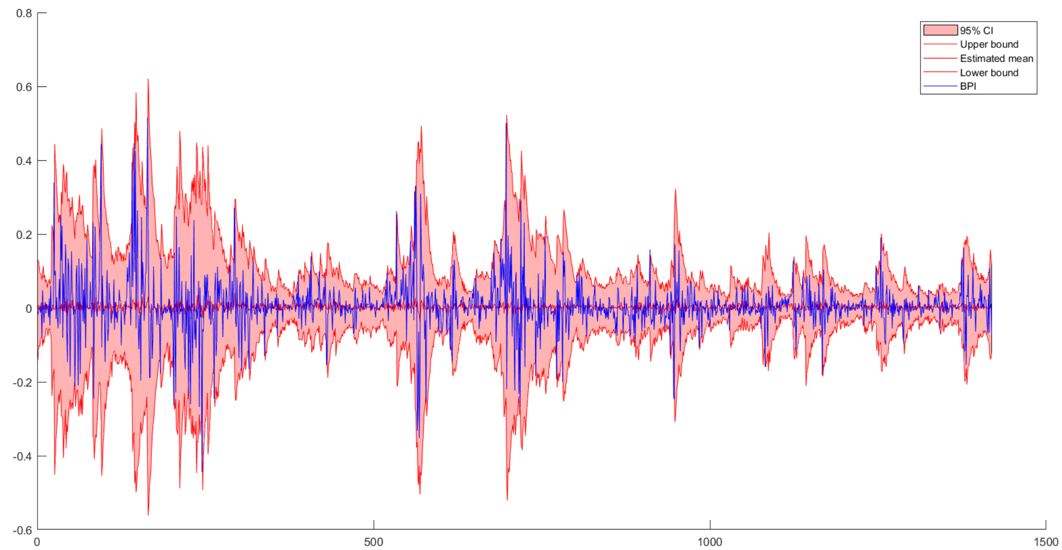

Figure 10: Log returns of BPI vs the estimated 95% CI using an ARX(1)-GARCHX(1,1)

ure 10. We saw that now most of the realizations actually fall within the 95% confidence interval. This clearly shows that the calibrated model is working well and describes precisely the variance dynamics of the process over time. Moreover we can observe that the GARCH parameters are both significant, meaning that the past variance and shocks are still fundamental to estimate next variance values, even after introducing the several exogenous variables. Also, the GARCH term γ is noticeably bigger than the ARCH parameter α. This result is in accordance with the paper, and shows that past volatility effects are superior to past shock effects, indicating the phenomenon of variance persistence. On the other hand the punctual prediction is quite bad, always near to zero, as a proof of the low predictability of the series as expected in accordance with the no arbitrage principle and the high pValues of the test against the null hypothesis of parameter equal zero.

Then we decided to perform a variable selection to better tailor our model to our aim, and easing out the convergence of the minimization algorithm. We reduce the exogenous variable of the mean equation to only 5, while we still keep 10 out of the 16 initial regressors for the variance equation. To perform this variable selection we performed a group lieklihood ratio test. The relevance of this group of variables score a pValue of 0.56, definetely allowing us to eliminate them. In table 7 you can see the spared variable and the survived coefficients and their relevance.

The only variable with a pValue higher than 0.05 is the autoregressive term with $pValue = 0.07$, but considering that it is still next to the acceptance threshold we decided to keep it in the model to preserve the model structure proposed in the paper. Comparing this restricted model with the complete one we retrieve lower AIC and BIC, as shown in Table 8. The high quality of the restricted model is confirmed also by the

Parameter	Estimated value	Estimated pVal
$Constant\ \beta_0$	0.0028***	0.0046
$\Delta lnGold_{t-1}$	0.1962**	0.0379
$\Delta lnSSI_{t-1}$	-0.1629**	0.0132
$\Delta lnStoxx_{t-1}$	0.2433***	0.0015
$\Delta ChinaB_{t-1}$	-0.0585**	0.0106
$AR\ \delta$	0.0592*	0.0703
$Constant\ \lambda_0$	-19.5306***	0
$\Delta lnCNY_{t-1}$	-209.2611***	0.0005
$\Delta lnEUR_{t-1}$	193.2449***	0.0000
$\Delta lnGold_{t-1}$	-28.7075**	0.0247
$ChinaB_{t-1}$	4.2623***	0.0000
USB_{t-1}	-1.9242***	0.0000
$GerB_{t-1}$	2.8360***	0.0000
$JapB_{t-1}$	-5.3260***	0.0000
$ChinaI_{t-1}$	-0.8405***	0.0001
USI_{t-1}	4.0343***	0.0000
$GerI_{t-1}$	1.4891***	0.0014
$GARCH\ \gamma$	0.7711 ***	0
$ARCH\ \alpha$	0.1969***	0

Table 7: Mean parameters above, variance parameters below

plot of the 95% confidence interval, which doesn't lose any precision nor information w.r.t. the one of the complete model. This result supports the usage of a reduced model.
In both models the test for ARCH effects gives a pValue slightly over 1%. This is a noticeable improvement with respect to the pValue around 0 obtained with the ARMAX-GARCH model, although it still leave place for eventual further improvements, for example with higher order GARCH models. Studying the normalized residuals, we can observe them in Figure 11. They appear satisfyingly heteroskedastic.

Finally we go back to the other calibrated model, the ARMAX(4,2)-GARCHX(1,1). This model was kept in consideration as the ARMAX(4,2) was the best performing in the ARMAX-GARCH model selection previously performed. In order to avoid convergence problem, we exploit the solutions found from the calibration of ARMAX(4,2)-GARCH(1,1) and from the calibration of ARX(1)-GARCHX(1,1) for what concern the exogenous part of the variance equation. Also in this case the plot shows extremely satisfactory results from the point of view of the variance. Comparing the AIC and BIC indicators of this model with the ones of the complete ARX(1)-GARCHX(1,1) we observe that the higher number of parameters is justified by a better fit of the data

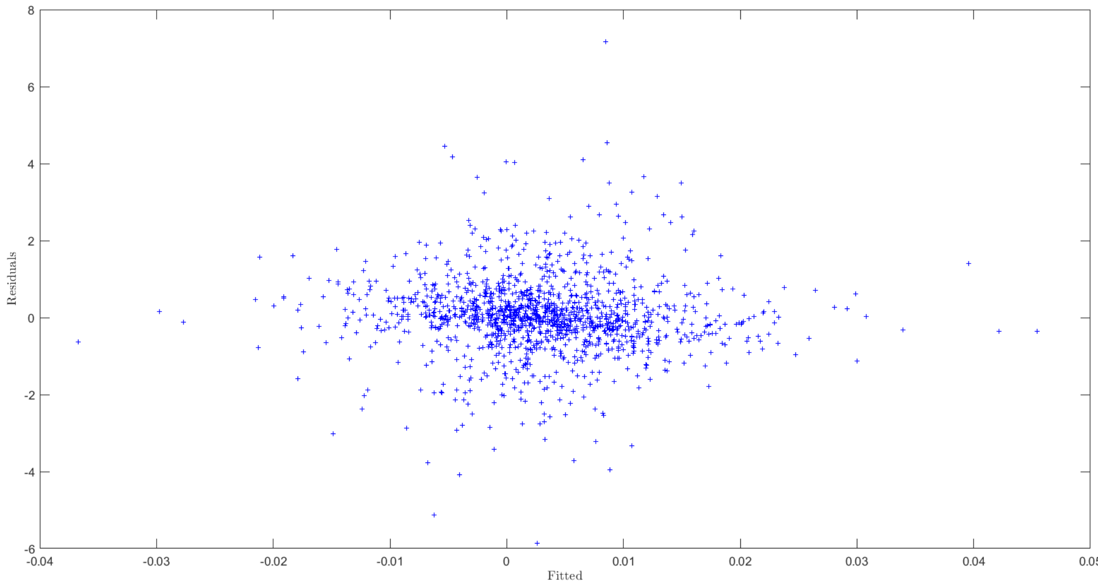

Figure 11: Normalized residuals with heteroskedastic behaviour

(Table 8).

Model	AIC	BIC
Complete	-4268	-4074
Restricted	-4288	-4188
$ARMAX(4,2)$	-4331	-4110

Table 8: AIC and BIC comparison of Complete, Restricted and ARMAX(4,2) models

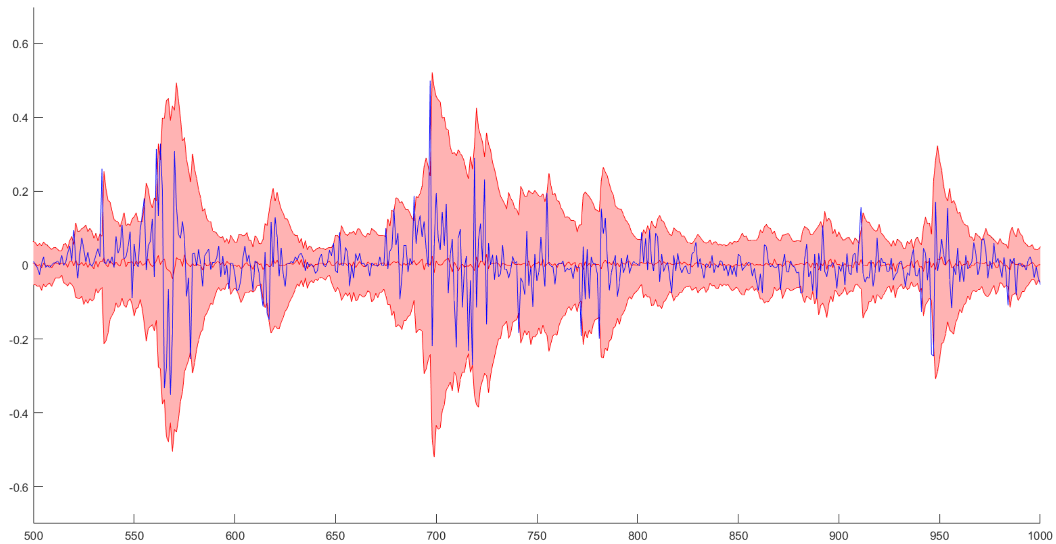

Figure 12: Zoom of ARX(1)-GARCHX(1,1)

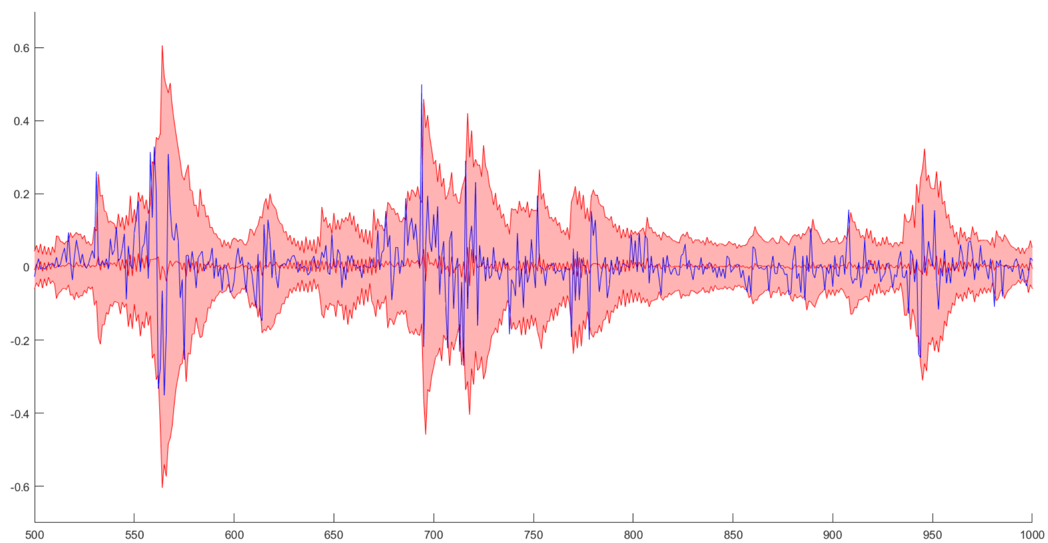

Figure 13: Zoom of ARMAX(4,2)-GARCHX(1,1)n

On the other hand, the test on the ARCH effects obtain a pValue equal to 0, showing that the model is unable to fit ARCH effects, in opposition to the over 1% obtained before. Due to the unsatisfying results on ARCH effects test we also attempted to implement an ARMAX(4,2)-GARCH(2,1), thus extending the ARCH part to 2 time steps. This model was however unable to improve the ARCH test results, and for this reason we didn't study this model more in detail.

In conclusion we try to answer the initial question: can Bitcoin become a viable alternative to fiat currencies? Looking at our analyses it can be seen that the volatility had a decreasing trend over the studied period and it could be meaningful to suppose a trend in direction of a more stable Bitcoin price. Also the volatility showed to be strongly related to several macroeconomic variables, especially with the Chinese one, as expected due to the higher trading volumes. In despite of this two signals, 4 years later, we can say that surely the decreasing trend of Bitcoin volatility didn't last long, and nowadays Bitcoin is an extremely volatile asset. On the other hand, it should be admitted that Bitcoin is always more accepted as a paying method, although its high volatility, keeping still open the possible path of Bitcoin to become a widely used currency in everyday life.

The analyses conducted in this project on different models let space to further development, on one side considering more complex structures, to better fit the ARCH effects, still present in the model. On the other hand it would be interesting to study the model calibrated on more recent data, in order to study more deeply how the situation evolved from 2017 to 2021, if the used model could still describe properly the series and eventually looking for new macroeconomic variable to add, especially in the variance equation, in accordance to nowadays trading volumes in different countries.