Assessing the Behavioural effect of Bitcoin price on the volumes: an RDD approach

Rosario Martucci

Claudio Pietro Pasina

Benjamin Bernedo Rendel

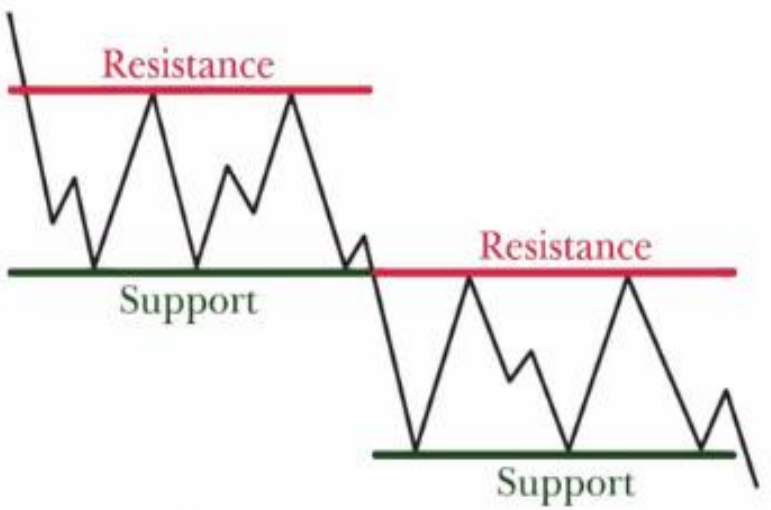
Thomas Capelletti

The Idea

Bitcoin is the main currency in the Crypto market, moving high volumes of exchanges due to its price volatility and the lack in regulamentations that make this kind of market more appealing for inverstors.

Bitcoin presents its own particular resistances or supports in the prices when studied with technical analysis.

Research Question: obtain the Behavioural effect on the volumes at a given threshold of the price of Bitcoin.



Paper References

Kjærland, Frode, Aras Khazal, Erlend A. Krogstad, Frans B. G. Nordstrøm, and Are Oust. 2018. "An Analysis of Bitcoin's Price Dynamics" *Journal of Risk and Financial Management* 11, no. 4: 63. https://doi.org/10.3390/jrfm11040063

The main objective of this study is to deepen the understanding of **the determinants that influence Bitcoin price fluctuations**, with the intention of providing investors with a robust information framework **for the formulation of informed investment decisions.**

The research involves factors of different types, related to Bitcoin such as the **Hashrate**, but also variables such as the returns of the **S&P 500** and **Google searches**.

The Data

- **BTC Data**: Bitcoin historical data, including price and volume, from **January 1, 2020, to April 30, 2024**. Source: **Yahoo Finance.**
- **S&P 500 Data** : S&P 500 close price from January 1, 2020, to April 30, 2024. **Missing values** have been interpolated using **local regression**. Source: **Yahoo Finance.**
- **Inflation Data**: The Data is collected on a year basis and it is the worldwide average inflation. Source: **International Monetary Fund (IMF)**.
- **US Interest Rate Data**: Monthly Federal Funds Rate from January 1, 2020, to April 30, 2024, repeated for each day. Source: **Federal Reserve Economic Data (FRED)**
- **Volatility Data**: Historical data for the **Crypto Volatility Index (CVI)**. Source: **Investing.com**

The Data

- **Market Capitalization Data**: gathered on monthly basis and adjusted for years. **Source: Statista.com**

- **Shock Dummies**: Dummy variables indicating significant events related to **Elon Musk and Tesla's actions regarding Bitcoin** have been created, covering specific dates of interest:

  - **February 8, 2021** Tesla announces a **$1.5 billion investment in Bitcoin** and plans to accept it as a form of payment for its products.
  - **May 16, 2021** Elon Musk expresses **concerns about Bitcoin's environmental impact** on Twitter, leading to a significant drop in its price.
  - **July 20, 2022** Tesla announces it **sold 75% of its Bitcoin holdings** in Q2 2022.

Data Merging: All datasets have been **merged based on the common Date column**.

The RDD Model

$$BTC\_Volume_t = \alpha_0 + \beta_1 D_t + \beta_2 (BTC\_Price_t - THD) + \beta_3 X_t + e_t$$

Where:

- **BTC_Volume**: Dependent variable, representing the volume of Bitcoin traded.
- **D**: Dummy variable indicating whether the Bitcoin price is above or below the threshold.
- **(BTC_Price - threshold)**: represents the difference between the Bitcoin price and the specified threshold.
- X_t:
  - **SP500**: S&P 500 index.
  - ***Inflation***: Inflation rate.
  - **IntRateUS**: Interest rate in the US.
  - ***CVI***: Crypto Volatility Index.
  - **BTC_MKTCAP**: Bitcoin market capitalization.
  - **ELON_2_21**: Dummy event related to Elon Musk on February 8, 2021.
  - **ELON_5_21**: Dummy event related to Elon Musk on May 16, 2021.
  - **ELON_7_22**: Dummy event related to Elon Musk on July 20, 2022.

The RDD Model

Other **possible important controls** for which the data could not be freely gathered could be:

- **Sentiment**: reflects the **overall mood or attitude of investors towards Bitcoin**.
- **Subscriptions and active accounts on cryptocurrency exchange platforms**: this could provide valuable insights into the level of **investor participation in the Bitcoin market**.
- **Bitcoin hashrate**: **computational power expended by miners** to secure and maintain the Bitcoin network, could also influence trading volume.

Statistics

Variable	Minimum	Mean	Median	Maximum	SD
SP500	2237.4	4049.39	4124.58	5254.35	566.94
Inflation	3.2	5.85	5.9	8.7	2
IntRateUS	0.05	2.06	0.71	5.33	2.23
CVI	34.5	77.48	75.22	170.55	22.28
BTC_MKTCAP	117.81	604.28	566.66	1403.11	324.11
BTC_Price	4970.79	31012.07	28844.33	73083.5	16576.44
BTC_Volume	5331172801	32259117032.66	29204071162	350967941479	18574174866.17

Table of the main statistics for each variable in the model

Correlation

SP500 and MKTCAP are very highly correlated with the X, creating **multicollinearity** problems.

Correlations with BTC_Volume (Y) are equilibrated and not problematic.

Correlation between BTC_Price and BTC_Volume are maybe a little too low.

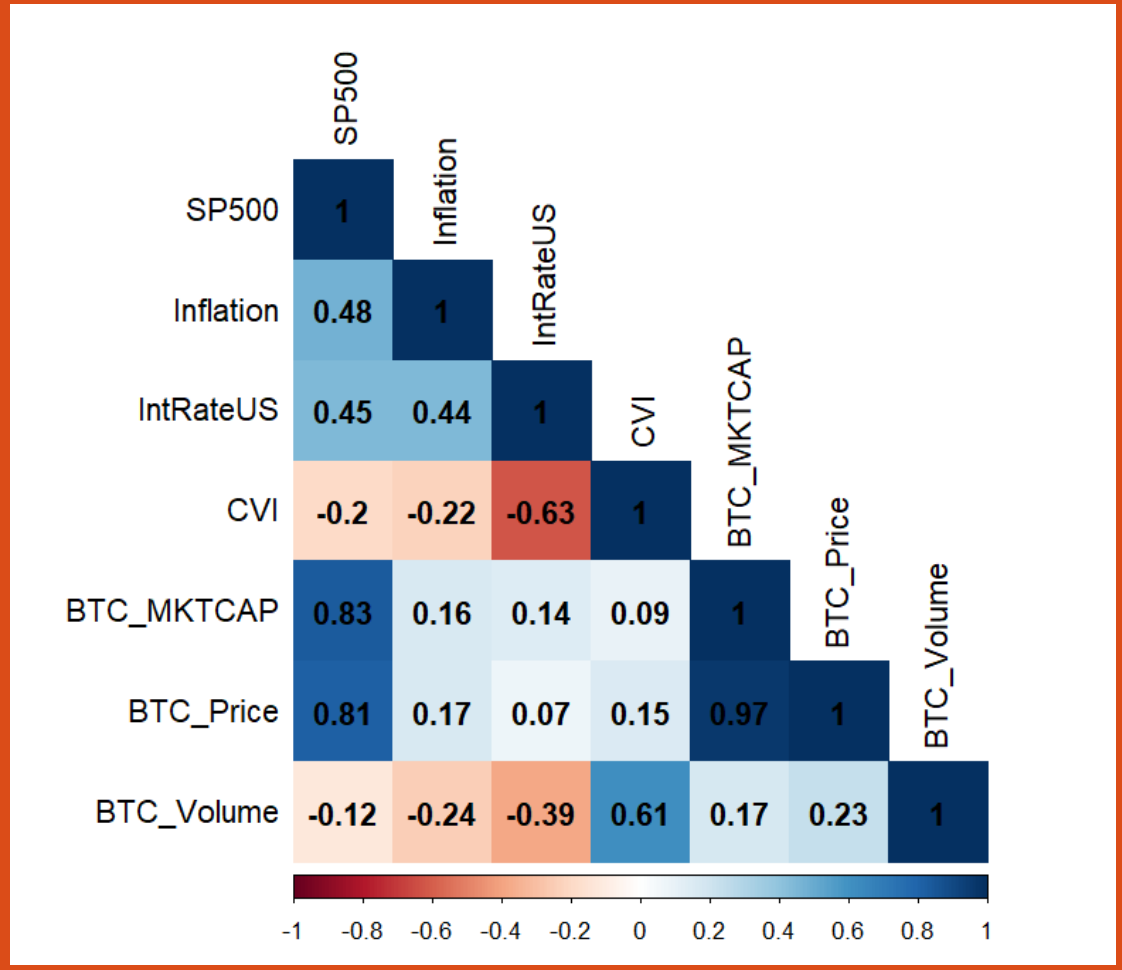


Table of the correlations between each variable

Threshold choice

For the choice of the threshold we took a value that was often mentioned on investing websites as an **important resistance/support** for BTC price **around 29000 BTC/USD.**

Moreover, the value was also a central value between the ones in the analysis period, allowing for more accurate estimates.

Further analysis made us adjust the discontinuity value to **29050 BTC/USD**.

Bitcoin Price in time.
The red line is the 29050 BTC/USD threshold.

RDD Assumptions and Tests

- **Continuity of the potential outcome**

- **CIA (Conditional Indipendance Assumption):**

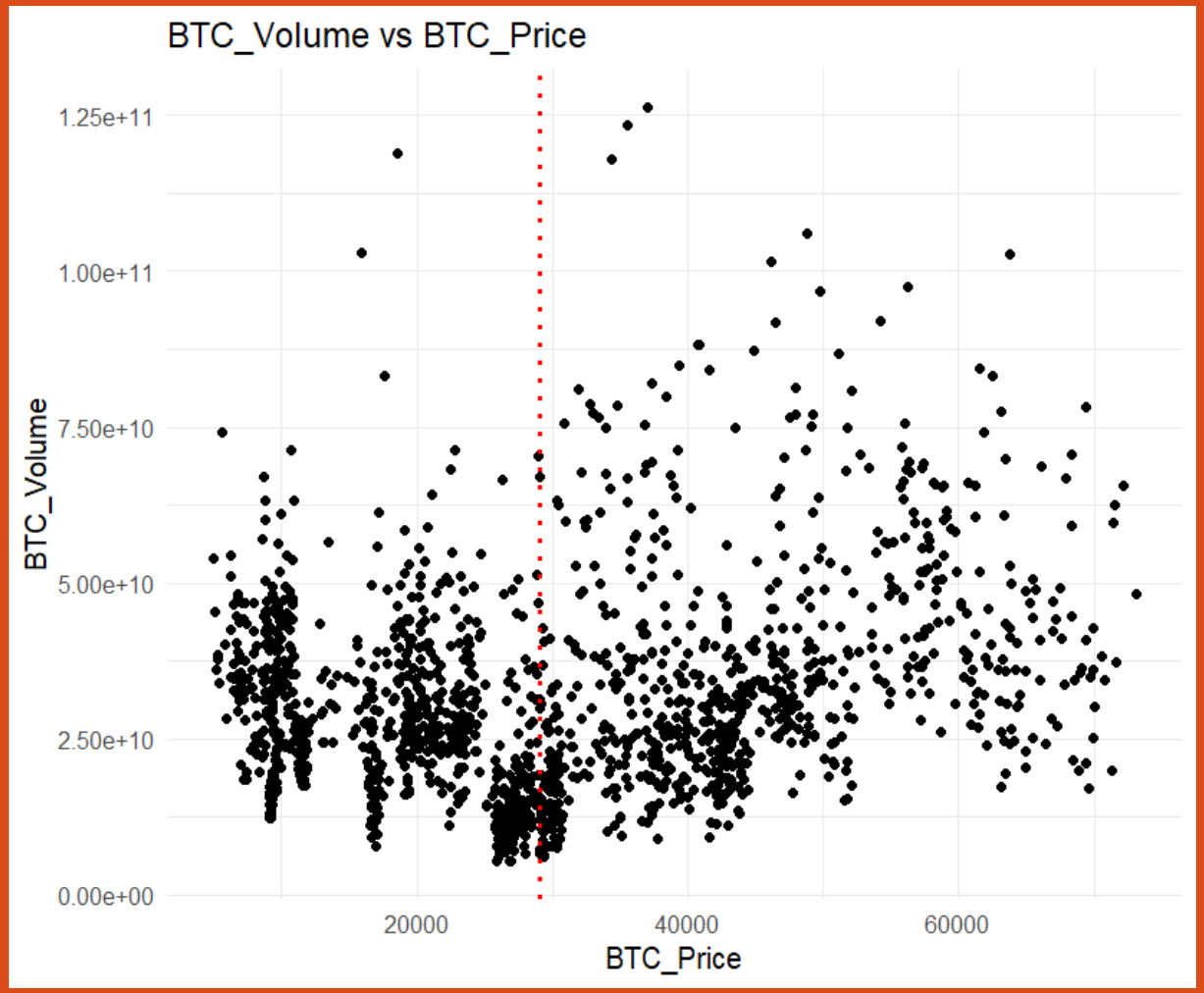
  - Discontinuity at the threshold
  - Continuity of the covariates
  - No perfect-manipulation of the assignment variable (McCrary Density test)
  - Bandwidth sensitivity

- **Additional binwidth test**

Outcome Continuity

We interpreted graphically the continuity of the outcome at the threshold.

From the scatterplot, there are **no particular differences before and after the threshold** in the volumes of Bitcoin.



Bitcoin Volume compared with Price in a scatterplot.
The red line is the selected Threshold of 29050 BTC/USD.

Discontinuity Test

The table summarizes the **rdrobust results** at three different threshold values.
Notably, at the **threshold of 29050 and 29100**, both the Conventional and Robust methods show **statistically significant treatment effects**, indicating that these threshold may be the most effective for detecting meaningful changes in the outcome variable (BTC_Volume).

H0: continuity
H1: discontinuity

Method	Threshold	Coefficient	Std_Error	z_value	P_value	Confidence_Interval
Conventional	29000	7849357629	4894062299	-1.604	0.109	[-17441543472.615, 1742828214.111]
Robust	29000	NA	NA	-1.868	0.062	[-20336345042.840, 490632977.620]
Conventional	29050	-11277128245	5269267959	-2.140	0.032	[-21604703669.095 , -949552820.045]
Robust	29050	NA	NA	-2.317	0.021	[-24229681518.861 , -2021773568.826]
Conventional	29100	-11972849840	5040185071	-2.375	0.018	[-21851431054.031 , -2094268626.083]
Robust	29100	NA	NA	-2.550	0.011	[-24364529466.816 , -3187441137.057]

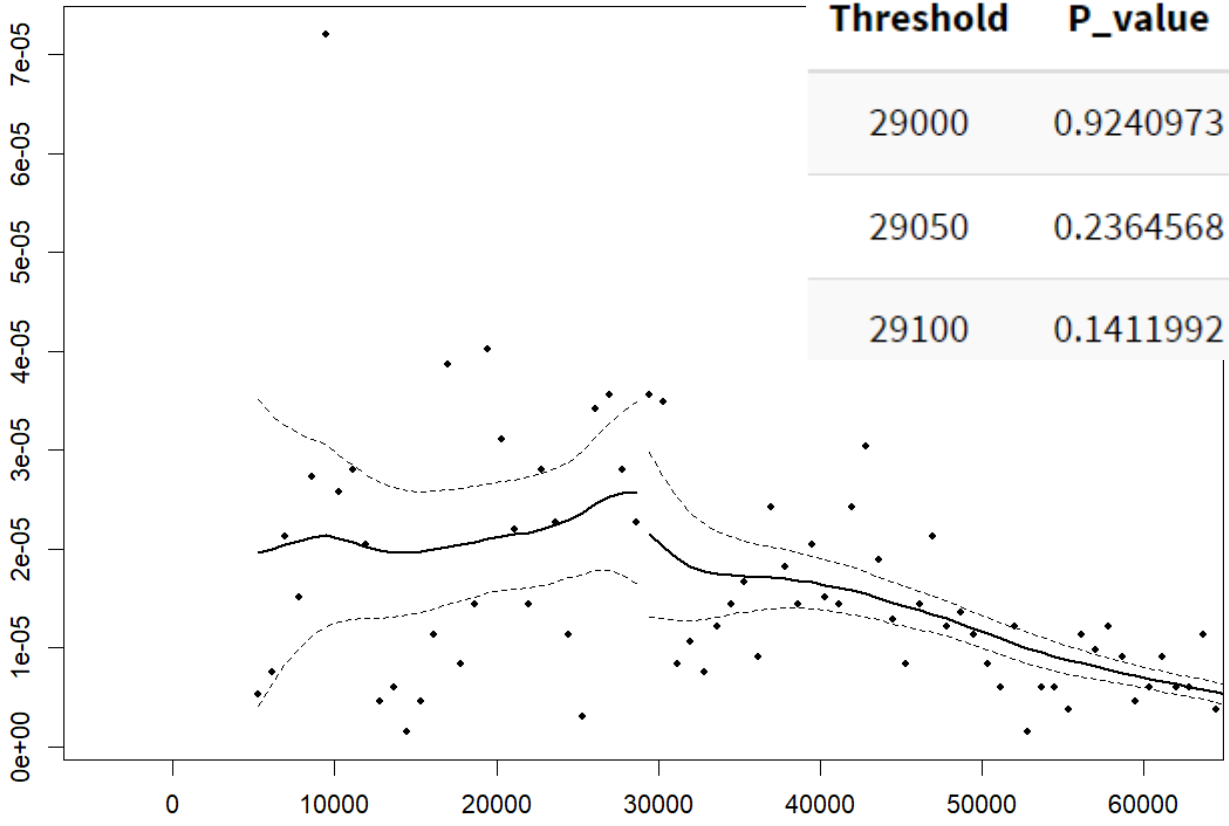
Table of the test performed with different thresholds around the chosen value

McCrary Test

The table summarizes the **McCrary test for density linearity** at three different threshold values.
A high p-value suggests a less statistically significant difference in the densities before and after the threshold.
At all the thresholds the values accept the H0 hypotesis of continuity.
Because of the **previous test** being **not significant at 29000** and because of the **quasi discontinuity at 29100**, we **model** the **RDD with the 29050 cutoff**.

H0: density continuity
H1: density discontinuity

Threshold	P_value
29000	0.9240973
29050	0.2364568
29100	0.1411992

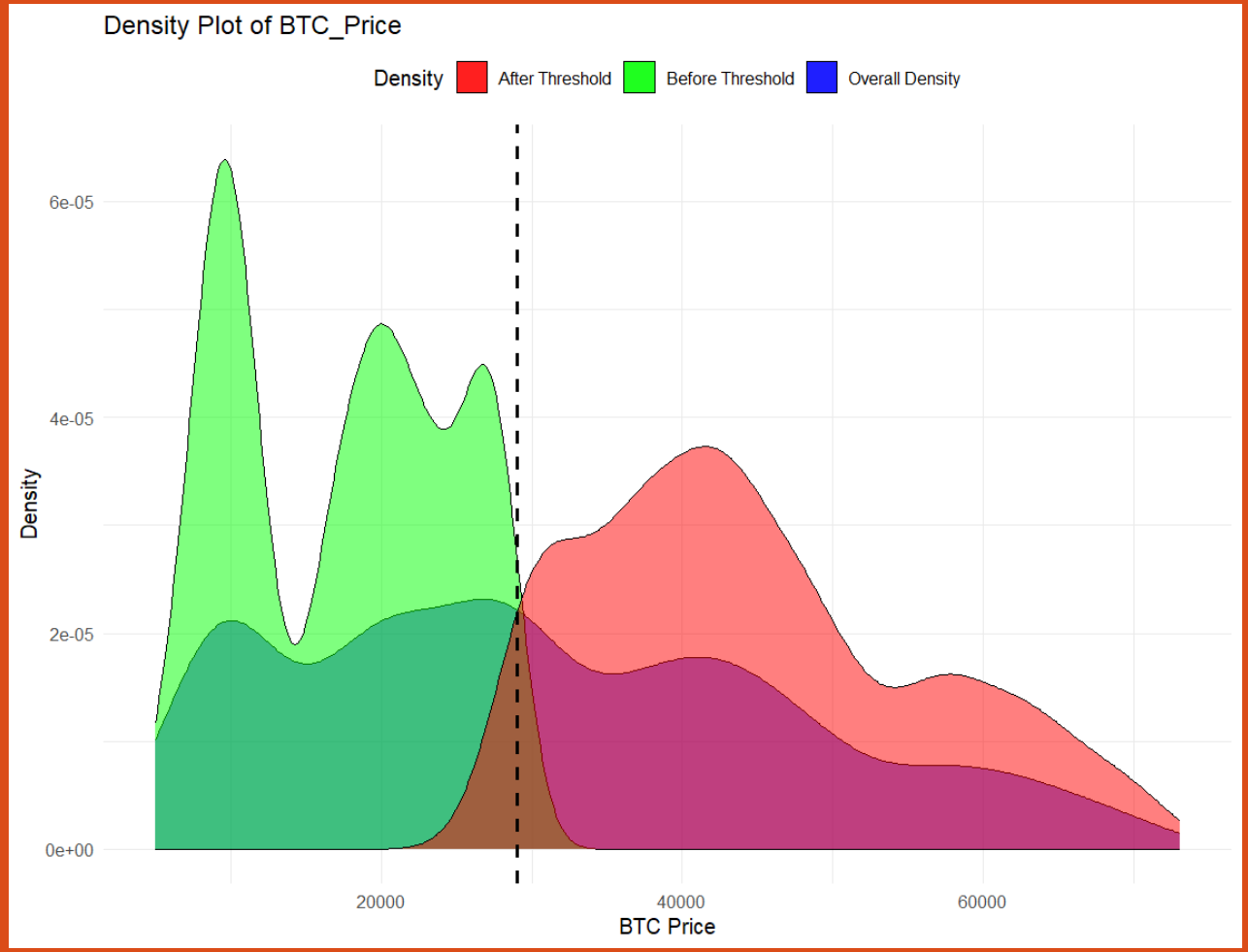


McCrary test graph for threshold of 29050 selected according to P-Values in the table and previous tests

Distribution analysis

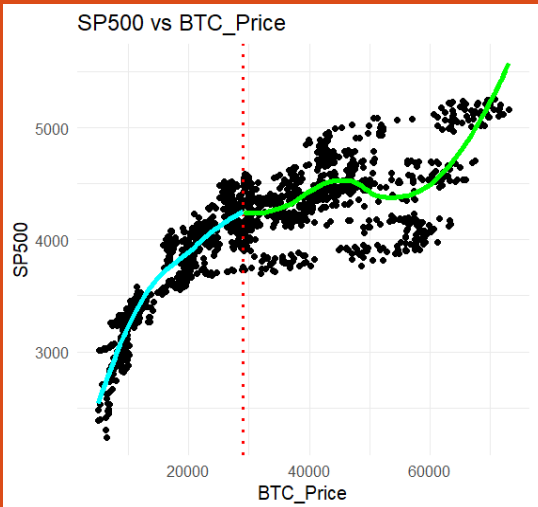
From the visual analysis of the density distributions we can see that the two separed distributions before and after the threshold **meet each other** perfectly at the **same height** of the overall distribution.

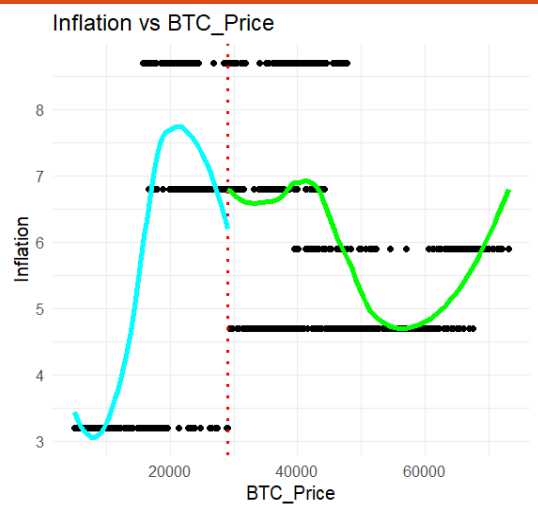
This suggests a **good continuity** at the threshold.

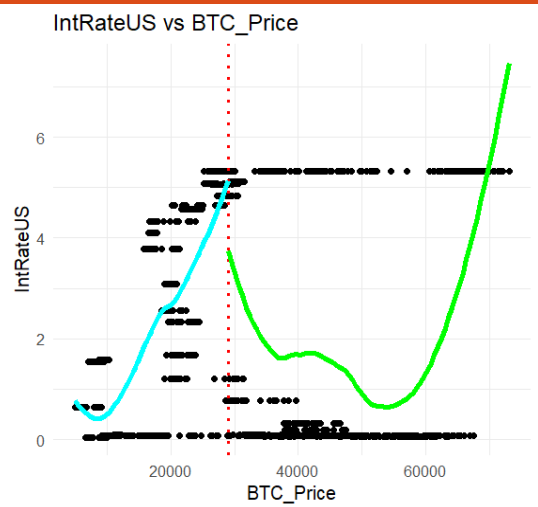


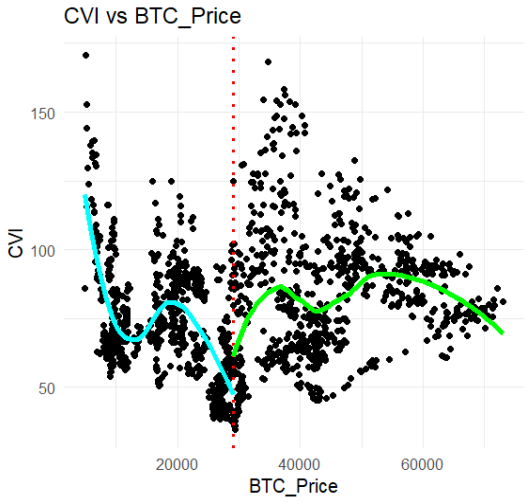
Density functions analysis around threshold

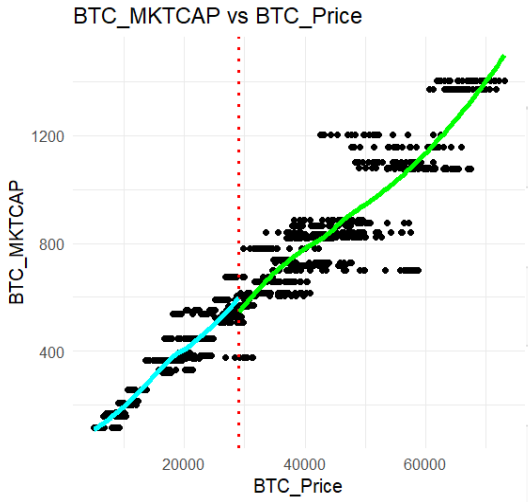
Continuity of the controls at the threshold











Variables	P_value	Robust_P_value
SP500	0.001	0.000
Inflation	0.187	0.359
IntRateUS	0.324	0.184
CVI	0.032	0.037
BTC_MKTCAP	0.000	0.000

H0: continuity
H1: discontinuity

Continuity tests.
The test shows that the variables SP500 and BTC_MKTCAP are discontinuous, even if graphically the interpretation seems wrong: the visual inspection might not detect small differences which are statistically significant (scale).
We keep CVI even if it is slightly discontinuous for its importance in the models.

Continuity of the controls at the threshold

Variable	Bandwidth	p_value	Mean_Group_0	Mean_Group_1
CVI	Full Sample	< 2.2e-16	72.673640	82.345760
Inflation	Full Sample	3.546e-05	5.647935	6.062276
IntRateUS	Full Sample	0.0006726	2.253066	1.870524
SP500	Full Sample	< 2.2e-16	3673.738000	4433.522000
BTC_MKTCAP	Full Sample	< 2.2e-16	351.644000	862.104600
CVI	500	0.05719	70.214450	54.438590
Inflation	500	0.656	6.943750	7.150000
IntRateUS	500	0.2032	3.151250	4.043438
SP500	500	0.004148	4093.040000	4291.836000
BTC_MKTCAP	500	0.463	561.048100	549.575300

Control ~ D
H0: difference in means is 0
H1: difference in means isn't 0

An additional **T-test** for the continuity was performed to see if the predicted averages before and after a cutoff represented by a dummy D were significantly different.

The test was repeated with full estimated values and with a smaller local bandwidth.

Local results proved that SP500 was the only variable with a significant discontinuity.

Parametric RDD Models Results

The table shows the results for the main parametric models used for the analysis.

Model 3 is the most favorable choice due to its robust coefficients and strong model fit. **The estimated behavioral effect indicates a significant loss in volumes of 5.33 billion as the threshold is crossed**.

The models with the S&P500, Interest Rates and the Marketcap are not reported because these variables were stealing the variability from the D, making the value non relevant. Only **one dummy** improved the model but only if no other variables were added.

	(1)	(2)	(3)	(4)	(5)
(Intercept)	4.164126e+15***	8.356507e+15***	6.064331e+15***	6.400199e+15***	6.355224e+15***
	(8.385492e+14)	(7.962224e+14)	(6.293205e+14)	(6.491155e+14)	(6.701848e+14)
I(poly(BTC_Price, 1) - 29050)	1.433423e+11***	2.876580e+11***	2.087547e+11***		
	(2.886570e+10)	(2.740867e+10)	(2.166334e+10)		
D	1.243553e+09	-6.965352e+09***	-5.330606e+09***	-4.203553e+09***	-4.354280e+09**
	(1.452016e+09)	(1.394652e+09)	(1.092781e+09)	(1.219011e+09)	(1.340252e+09)
ELON_7_22		-1.455267e+10***			
		(7.981283e+08)			
Inflation			-1.366487e+09***	-1.138158e+09***	-1.116530e+09***
			(1.587457e+08)	(1.929430e+08)	(2.088534e+08)
CVI			4.570890e+08***	4.554679e+08***	4.550357e+08***
			(1.441212e+07)	(1.441811e+07)	(1.451029e+07)
I(poly(BTC_Price, 2) - 29050)1				1.873699e+11***	
				(2.396405e+10)	
I(poly(BTC_Price, 2) - 29050)2				3.294653e+10*	
				(1.585909e+10)	
I(poly(BTC_Price, 3) - 29050)1					1.896119e+11***
					(2.535880e+10)
I(poly(BTC_Price, 3) - 29050)2					3.331670e+10*
					(1.592247e+10)
I(poly(BTC_Price, 3) - 29050)3					-4.160299e+09
					(1.535337e+10)
Num.Obs.	1581	1581	1581	1581	1581
R2	0.061	0.224	0.484	0.485	0.485
R2 Adj.	0.060	0.223	0.482	0.484	0.483
AIC	78835.0	78534.5	77892.9	77890.6	77892.5
BIC	78856.4	78561.4	77925.1	77928.2	77935.4
Log.Lik.	-39413.487	-39262.267	-38940.462	-38938.298	-38938.262
F	51.141	152.076	369.217	296.858	247.248
RMSE	16236510624.03	14755476684.90	1.2e+10	12021544749.61	12021264366.41

Binwidth testing

H0: D = 0
H1: D different from 0

The model provides **stable ATE** for all the significant models of the previous table.

The **estimates are slightly higher** in magnitude for the model with only the **dummy**.

Even if the model seems **slightly sensitive to binning**, it is coherent in the estimates.

Model	Bin Width	ATE (D)	MSE	P-Value
lm linear Musk dummy	100	-7902559137	1.513797e+20	0.0001886
lm linear w controls	100	-6202879313	1.117952e+20	0.0003901
lm quadratic w controls	100	-6847305854	1.117346e+20	0.0013030
lm cubic w controls	100	-7512513894	1.114902e+20	0.0007160
lm linear Musk dummy	500	-9809601936	6.796449e+19	0.0018585
lm linear w controls	500	-5737470761	4.766841e+19	0.0209408
lm quadratic w controls	500	-8758620192	4.683003e+19	0.0061449
lm cubic w controls	500	-8515220618	4.666193e+19	0.0081992
lm linear Musk dummy	1000	-10723847127	4.405583e+19	0.0045185
lm linear w controls	1000	-6179040446	2.722219e+19	0.0244684
lm quadratic w controls	1000	-7165738214	2.714691e+19	0.0499955
lm cubic w controls	1000	-6990006526	2.617593e+19	0.0535054
lm linear Musk dummy	2000	-9201123533	2.726249e+19	0.0378903
lm linear w controls	2000	-5356887605	1.294239e+19	0.0586662
lm quadratic w controls	2000	-6361729617	1.287358e+19	0.1015987
lm cubic w controls	2000	-6546770969	1.269294e+19	0.0970719

Local Regression RDD Results

H0: continuity
H1: discontinuity

Model	Estimate	SE	Robust Pr>|z|	Bandwidth	Rho (h/b)
No controls	-11.27 Bln	5.26 Bln	0.021	2560	0.420
Musk only	-9.62 Bln	2.98 Bln	0.001	2838	0.435
CVI only	-12.43 Bln	1.80 Bln	0.541	8452	0.602
Inflation only	-11.98 Bln	5.32 Bln	0.018	2478	0.435
Inflation + Musk	-10.60 Bln	5.44 Bln	0.037	2421	0.433

Model	Estimate	SE	Robust Pr>|z|	Bandwidth	Rho (h/b)
p(1)	-10.60 Bln	5.44 Bln	0.037	2421	0.433
p(2)	-12.71 Bln	5.74 Bln	0.023	4948	0.606
p(3)	-14.55 Bln	5.36 Bln	0.010	8799	0.777
p(4)	-15.50 Bln	5.35 Bln	0.004	13728	0.783

Table of the local regression results.

The selected controls are **ELON_7_22** and **Inflation** because of the multicollinearity issues of the other ones and the poor performances of CVI and IntRateUS.

The Bandwidth was chosen through cross validation by the rdrobust function.

The best performing model was then used to compare between different polynomials.

Polynomial models have **higher bias**.

Bandwidth Sensitivity

The model reported was the less sensitive polynomial. Without fixing Rho it would have been set to 1, building highly biased models.

The models presented show **significant sensitivity** to the bandwidth, making the **estimates** of the non parametric model **not so reliable**.

Also the magnitude of the estimates changes much with the choice of the bandwidth.

rdrobust(BTC_Volume , BTC_Price, c=29050, covs=(Inflation + ELON_7_22), p=(2))

H0: continuity
H1: discontinuity

Fixed	Estimate	SE	Robust Pr>|z|	Bandwidth
Rho=0.5	-12.71 Bln	5.74 Bln	0.022	4248
Rho=0.5 and h=1000	-12.92 Bln	14.00 Bln	0.367	1000
Rho=0.5 and h=2000	-15.00 Bln	9.82 Bln	0.137	2000
Rho=0.5 and h=5000	-12.82 Bln	5.68 Bln	0.020	5000
Rho=0.5 and h=10000	-4.32 Bln	3.33 Bln	0.145	10000

Table of the Bandwiths of the best performing model in terms of stablilty

Conclusions

Since we verified all the Assumptions, the value estimate of **–5.33 billion of Volume** in the linear model is considered a **behavioural** attributed **ATE** where the treatment is **the crossing of the cutoff price**. The estimates of the **non parametric models** were too biased to be considered as an ATE, even if they suggested that the **real ATE could be slightly more negative** than the one we estimated.

The model could be further improved by **adding additional controls** that could not be gathered for free on the web.

Overall the parametric models perform quite well while non parametric estimates were too sensitive to the bandwidth choice.

The model is **Internally valid** only at the threshold and can be repeated at different ones after checking the assumptions. The model is an **Externally valid** approach after for other currencies after the assumptions are verified for the selected threshold.

References

- https://bse2nse.com/fundamental-technical-analysis/10773-support-resistance-technical-analysis.html
- https://www.statista.com/statistics/377382/bitcoin-market-capitalization/
- https://it.investing.com/indices/crypto-volatility-index-historical-data
- https://fred.stlouisfed.org/series/FEDFUNDS
- https://www.imf.org/external/datamapper/PCPIPCH@WEO/OEMDC/ADVEC/WEOWORLD
- https://it.finance.yahoo.com/
- https://tradersunion.com/news/cryptocurrency-news/show/3796-bitcoin-price-forecast-btc-to-test-29000-resistance/
- https://www.mdpi.com/1911-8074/11/4/63
- https://royalsocietypublishing.org/doi/epdf/10.1098/rsif.2014.0623
- https://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2020-Dublin/papers/EFMA%202020_stage-1301_question-Full%20Paper_id-405.pdf
- https://royalsocietypublishing.org/doi/epdf/10.1098/rsos.191863

Thank you for your attention!

	(1)	(2)	(3)	(4)	(5)
(Intercept)	4.164126e+15***	8.356507e+15***	6.064331e+15***	6.400199e+15***	6.355224e+15***
	(8.385492e+14)	(7.962224e+14)	(6.293205e+14)	(6.491155e+14)	(6.701848e+14)
I(poly(BTC_Price, 1) - 29050)	1.433423e+11***	2.876580e+11***	2.087547e+11***		
	(2.886570e+10)	(2.740867e+10)	(2.166334e+10)		
D	1.243553e+09	-6.965352e+09***	-5.330606e+09***	-4.203553e+09***	-4.354280e+09**
	(1.452016e+09)	(1.394652e+09)	(1.092781e+09)	(1.219011e+09)	(1.340252e+09)
ELON_7_22		-1.455267e+10***			
		(7.981283e+08)			
Inflation			-1.366487e+09***	-1.138158e+09***	-1.116530e+09***
			(1.587457e+08)	(1.929430e+08)	(2.088534e+08)
CVI			4.570890e+08***	4.554679e+08***	4.550357e+08***
			(1.441212e+07)	(1.441811e+07)	(1.451029e+07)
I(poly(BTC_Price, 2) - 29050)1				1.873699e+11***	
				(2.396405e+10)	
I(poly(BTC_Price, 2) - 29050)2				3.294653e+10*	
				(1.585909e+10)	
I(poly(BTC_Price, 3) - 29050)1					1.896119e+11***
					(2.535880e+10)
I(poly(BTC_Price, 3) - 29050)2					3.331670e+10*
					(1.592247e+10)
I(poly(BTC_Price, 3) - 29050)3					-4.160299e+09
					(1.535337e+10)

	(1)	(2)	(3)	(4)	(5)
Num.Obs.	1581	1581	1581	1581	1581
R2	0.061	0.224	0.484	0.485	0.485
R2 Adj.	0.060	0.223	0.482	0.484	0.483
AIC	78835.0	78534.5	77892.9	77890.6	77892.5
BIC	78856.4	78561.4	77925.1	77928.2	77935.4
Log.Lik.	-39413.487	-39262.267	-38940.462	-38938.298	-38938.262
F	51.141	152.076	369.217	296.858	247.248
RMSE	16236510624.03	14755476684.90	1.2e+10	12021544749.61	12021264366.41

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001