Sentiment, Google Queries and Explosivity in the Cryptocurrency Markets

Pagnottoni Paolo

Department of Economics and Management

University of Pavia

3° Crypto Asset Lab Conference

University of Milan-Bicocca

5 November 2021

Background

- Since cryptocurrencies were conceived in first place under the advent of Bitcoin, research on this topic **has been prospering** in a highly multidisciplinary way
- The emergence of fields such as **machine learning (ML), deep learning (DL), Big Data Analytics (BDA), eXplainable artificial intelligence (XAI) and Automated Trading Systems (ATS)** have, at the same time, brought new challenges to both academics and practitioners calling for rapid knowledge advances in several disciplines
- On the other hand, methods pertaining to such innovative fields often **lack a sound and robust state-of-the-art econometric and statistical framework**, so to develop models accurately, conduct appropriate inference, and - most importantly - improve **model performances**
- It is against this background that we propose a **novel sentiment-based testing procedure for cryptocurrency explosiveness, which explicitly takes into account for any possible price predictors.**
- We develop our empirical application upon a **large set of financial news**, in order to anticipate speculative bubble occurrences in cryptocurrency prices.

Background

- The lack of fundamental values in the **cryptocurrency market** paves the way for the rise of unprecedented **speculative bubble phenomena**
- Recent works provide empirical evidence of the **presence of bubbles** in cryptocurrency prices – see e.g. Fry and Cheah (2016), Corbet and Yarovya (2018)
- We propose to exploit the information derived from a large set of cryptocurrency news to **detect** and, possibly, **anticipate** the presence of **speculative bubbles** in cryptocurrency prices
- By means of a **Covariate Superior Augmented Dickey-Fuller (CSADF) test**, we explicitly account for market sentiment when testing the presence of an explosive behaviour in cryptocurrency prices and we use the test results as an **early warning** tool
- Our results show that the covariate test statistics diverges significantly from the ADF test statistics in concomitance of price surges, highlighting its **ability to foresee speculative bubble occurrences**

Methodological problem & solutions

- From an econometric point of view, one of the main research questions related to cryptocurrencies concerns the possible **presence of bubbles** in their price
- An asset bubble is defined in literature as an **extreme price acceleration** that cannot be driven by the underlying fundamental economic variables (Case,Dreger)
- The end of this phase, often referred to as **bubble burst**, leads to drastic price drops, causing severe losses to investors and Several recent works provided **empirical evidence of the presence of bubbles in the cryptocurrency prices** (Fry,Corbet)
- Indeed, the extremely rapid price increase, to which the definition of financial bubble refers to**, can be described by an exponential growth, whose occurrence can be detected through right-sided unit root tests**
- Specifically, Phillips proposed a univariate approach to test for end-of-sample bubbles through estimation of the ADF regression

Methodology – CADF and (B)SCADF

- The **covariate-augmented ADF (CADF)** test by Hansen (1995) extends the traditional ADF test by checking the null hypothesis of unit root (φ=1) against the alternative of stationarity (φ<1) based on the following specification:

$$y_t = \mu + \phi y_{t-1} + \sum_{j=1}^{J} \psi_j \Delta y_{t-j} + \sum_{k=1}^{K} \xi_j \Delta x_{t-k} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

- **Right-sided ADF tests for unit root called Superior ADF (SADF)** tests were recently proposed in the literature (Phillips and Yu, 2011). In SADF tests, unit root is tested against the alternative hypothesis of *explosivity* (φ>1).

- We propose to merge these two approaches by using a **covariate-augmented (Backward) Superior ADF test ((B)CSADF),** based on the statistic

$$BSCADF_{r_2} := \sup_{r_1 \in [0, r_2 - r_0]} \left\{ CADF(p)_{r_1}^{r_2} \right\}$$

that is the supremum of the BSCADF test statistics computed over all possible end dates for samples starting at time t=1 (subject to a minimum sample size of r_0T).

Empirical analysis - target

- By repeating the ADF test and the CSADF test with various sentiment indicators as a covariate we evaluate whether sentiment is able to give a **signal of explosive behaviour** in the cryptocurrency prices that can act as an early warning for bubbles

- A further benefit of such a recursive procedure is that it allows us to select the optimal number of covariate lags to be considered at each time window according to standard information criteria

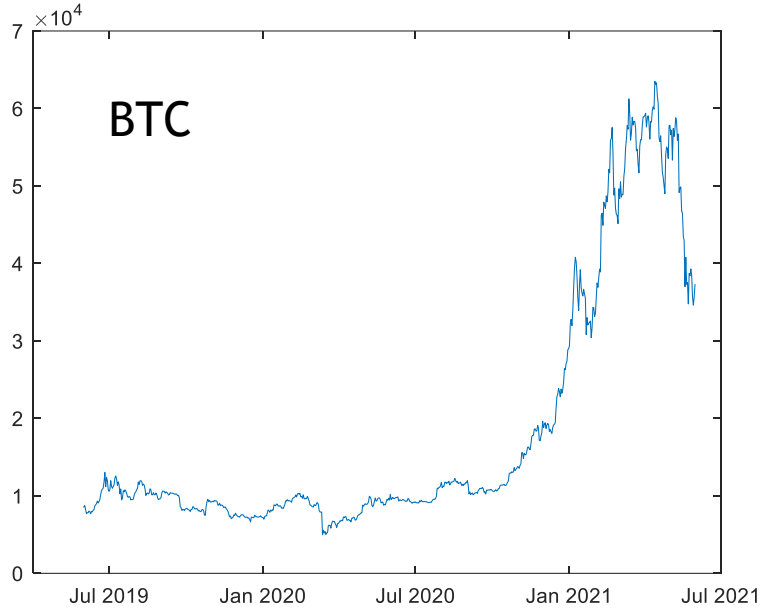
- If sentiment effectively reflects **market perceptions and expectation**, taking it into account can increase both the power and size performance of unit root tests

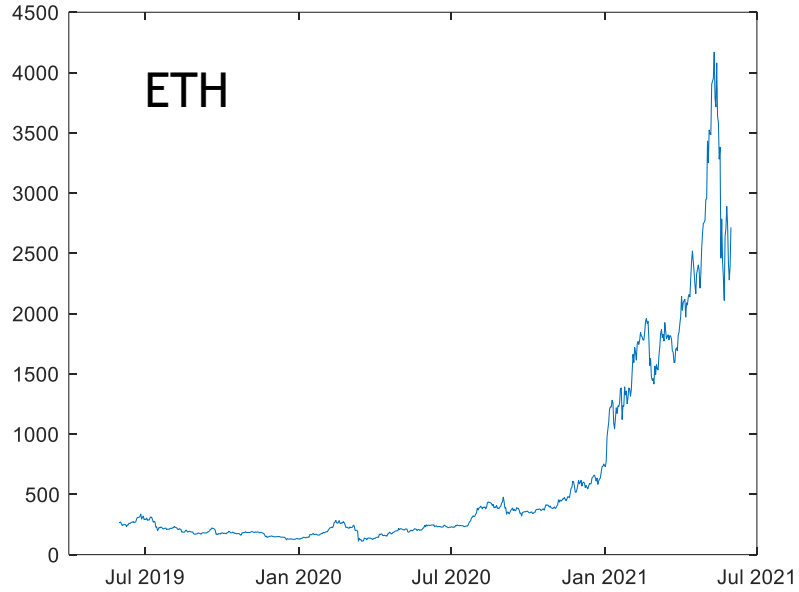
- **Bubble discovery:** which component dominates crypto-price explosive dynamics? Sentiment versus *volume*

Data

- **Cryptocurrency data:** daily prices (31 May 2019 – 31 May 2021)
  - Bitcoin (BTC)
  - Ethereum (ETH)
  - Ripple (XRP)
- **Exogenous covariates:** we consider two types of sentiment data
  - **Brain Sentiment data:** Sentiment and volume indicators, which monitor public financial news on cryptocurrencies from about 2000 financial media sources, as a result of a combination of various natural language processing techniques. The sentiment score assigned to each cryptocurrency is a value ranging from -1 (most negative) to +1 (most positive)
  - **Google Queries data:** a search trends feature that shows how frequently a given search term is entered into Google's search engine relative to the site's total search volume over a given period of time. Google Trends can be used for comparative keyword research and to discover event-triggered spikes in keyword search volume

Descriptive Statistics - Cryptocurrency data





Returns	Min	Max	Avg	Std	Skew	Kurt
BTC	-0.3717	0.1875	0.0028	0.0402	-0.8118	12.1680
ETH	-0.4235	0.2595	0.0046	0.0517	-0.6662	9.5027
XRP	-0.4233	0.5601	0.0034	0.0676	1.5363	16.7793

Descriptive Statistics – Brain Sentiment data

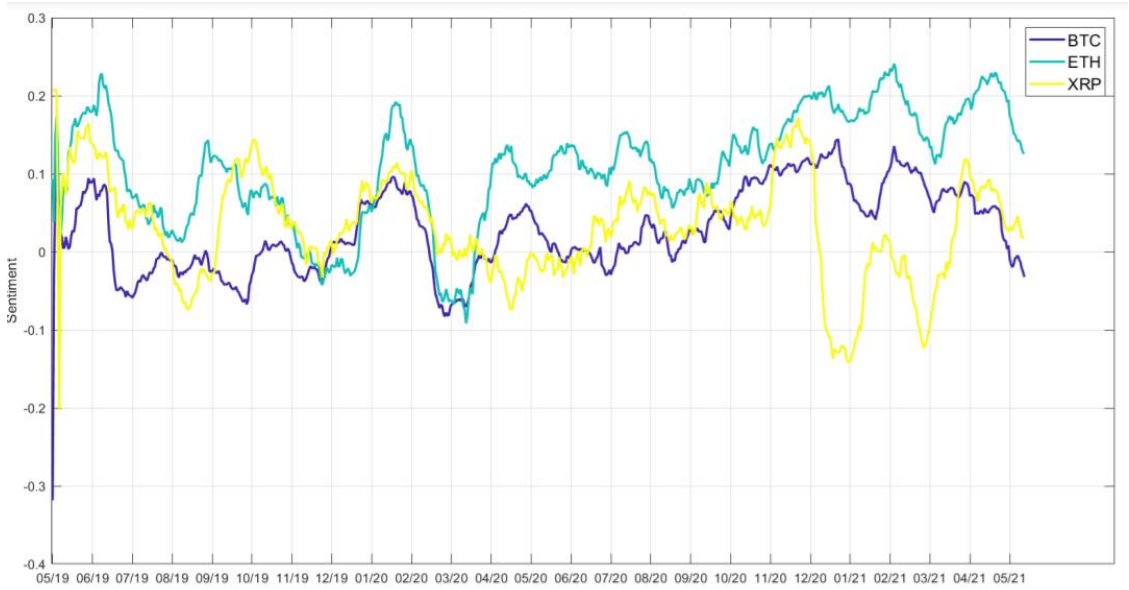


Figure 1: Brain Sentiment Indicator time series. The figure shows the dynamics of the daily sentiment index for the three selected cryptocurrencies over the period 31 May 2019 - 31 May 2021.

Brain Sentiment	Min	Max	Avg	Std	Skew	Kurt
BTC	-0.3182	0.1439	0.0283	0.0536	-0.2580	1.2452
ETH	-0.0916	0.2408	0.1083	0.0712	-0.4925	-0.2632
XRP	-0.2031	0.2072	0.0318	0.0655	-0.2622	0.1932

Descriptive Statistics – Google Queries data

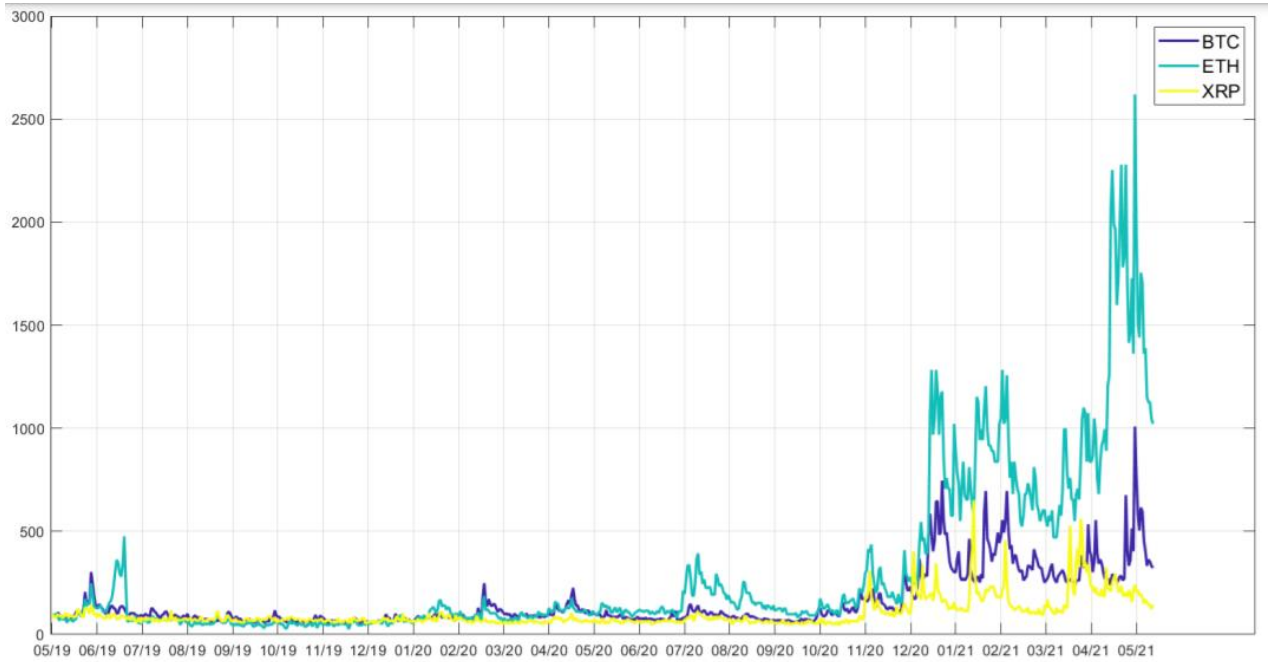


Figure 2: Google Search Index time series. The figure shows the dynamics of the daily Google Search Index for the three selected cryptocurrencies over the period 31 May 2019 - 31 May 2021.

Google Trends	Min	Max	Avg	Std	Skew	Kurt
BTC	52.2727	1008.8601	154.9287	131.3087	2.1707	5.4673
ETH	26.0870	2619.3434	296.0966	405.3924	2.4256	6.3983
XRP	44.3860	651.8219	101.2452	69.8259	3.1347	13.5140

Empirical results – Rolling regression

Figure 6: Sentiment Index rolling regression coefficients. The figure shows the estimated coefficients associated to the lagged Brain Sentiment index for the selected cryptocurrencies in a rolling linear regression exercise, where the response variable is the cryptocurrency return. The rolling window is set to 100 observations.

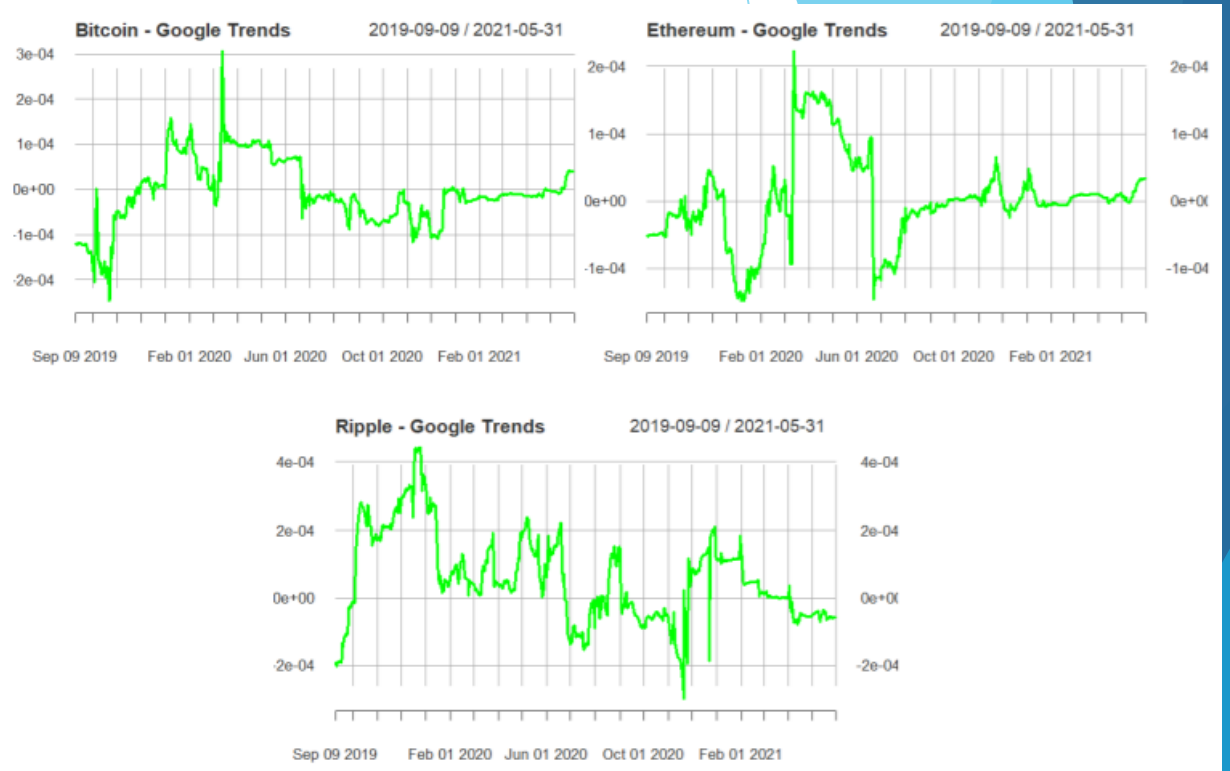


Figure 7: Google Search Index rolling regression coefficients. The figure shows the estimated coefficients associated to the lagged Google Search Index for the selected cryptocurrencies in a rolling linear regression exercise, where the response variable is the cryptocurrency return. The rolling window is set to 100 observations.

Empirical results – Brain Sentiment Index

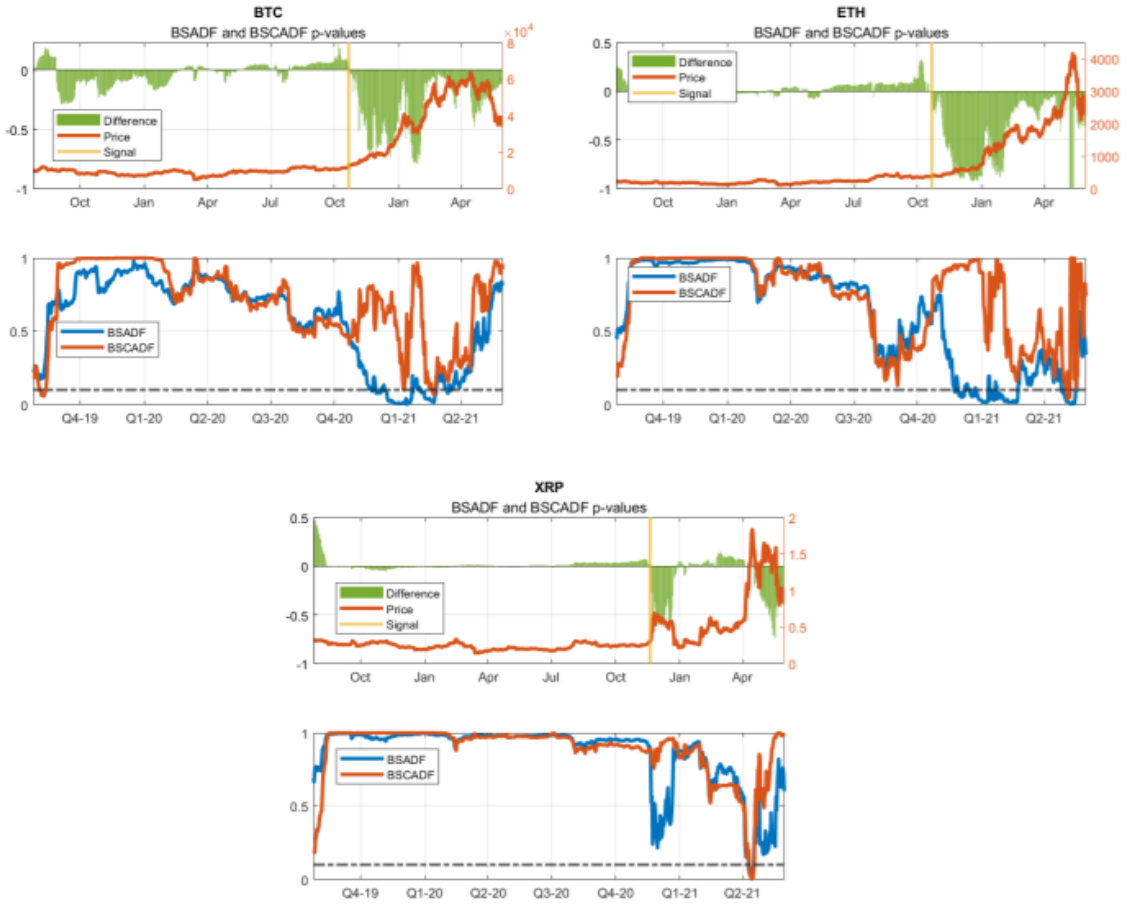


Figure 3: BSADF and BSCADF Brain Sentiment Index test p-values. The figure shows the BSADF and BSCADF p-values (bottom panels), their difference and closing price dynamics (top panels) of the three cryptocurrencies, using the Brain Sentiment Index as a covariate, over the full sample period. Black dashed lines indicate the 10% significance level.

Empirical results – Google Queries

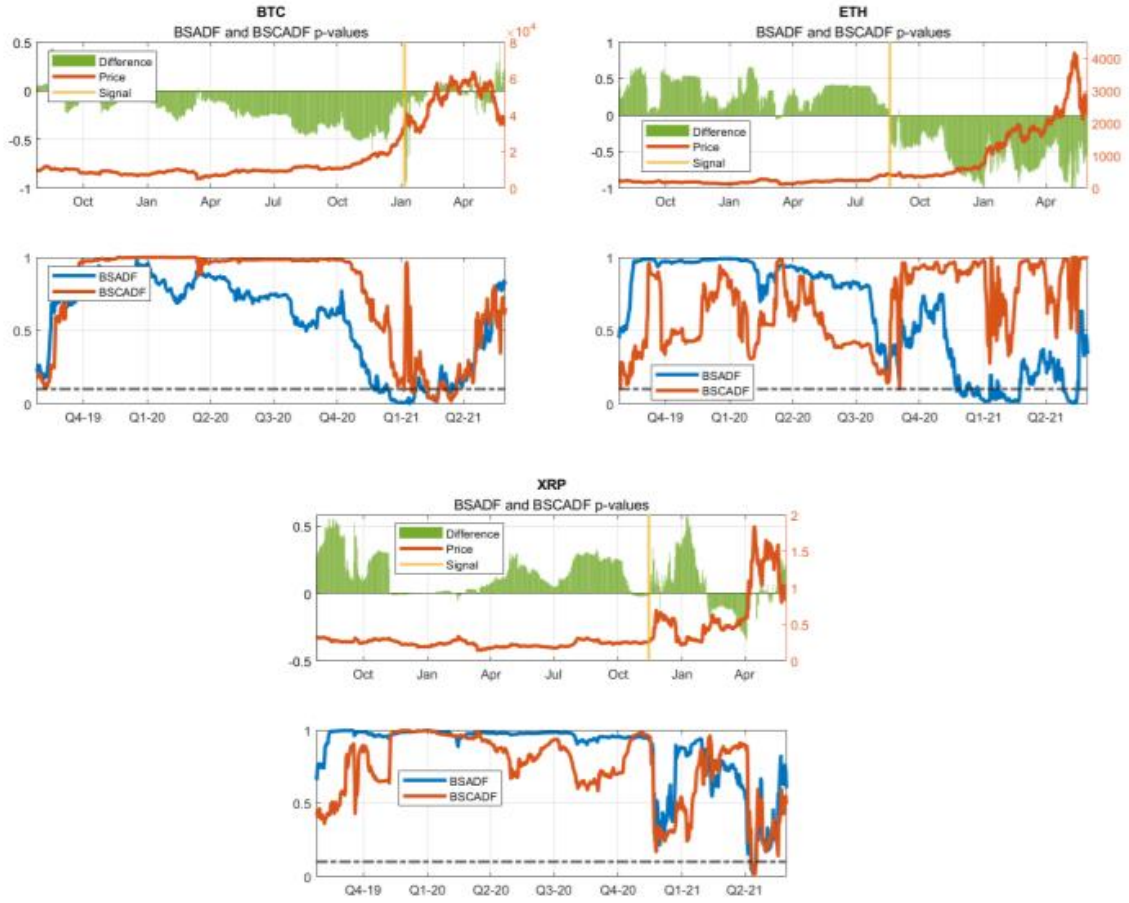


Figure 4: BSADF and BSCADF Google Search Index test p-values. The figure shows the BSADF and BSCADF p-values (bottom panels), their difference and closing price dynamics (top panels) of the three cryptocurrencies, using the Google Sentiment Index as a covariate, over the full sample period. Black dashed lines indicate the 10% significance level.

Empirical results – Relationship ADF/CADF

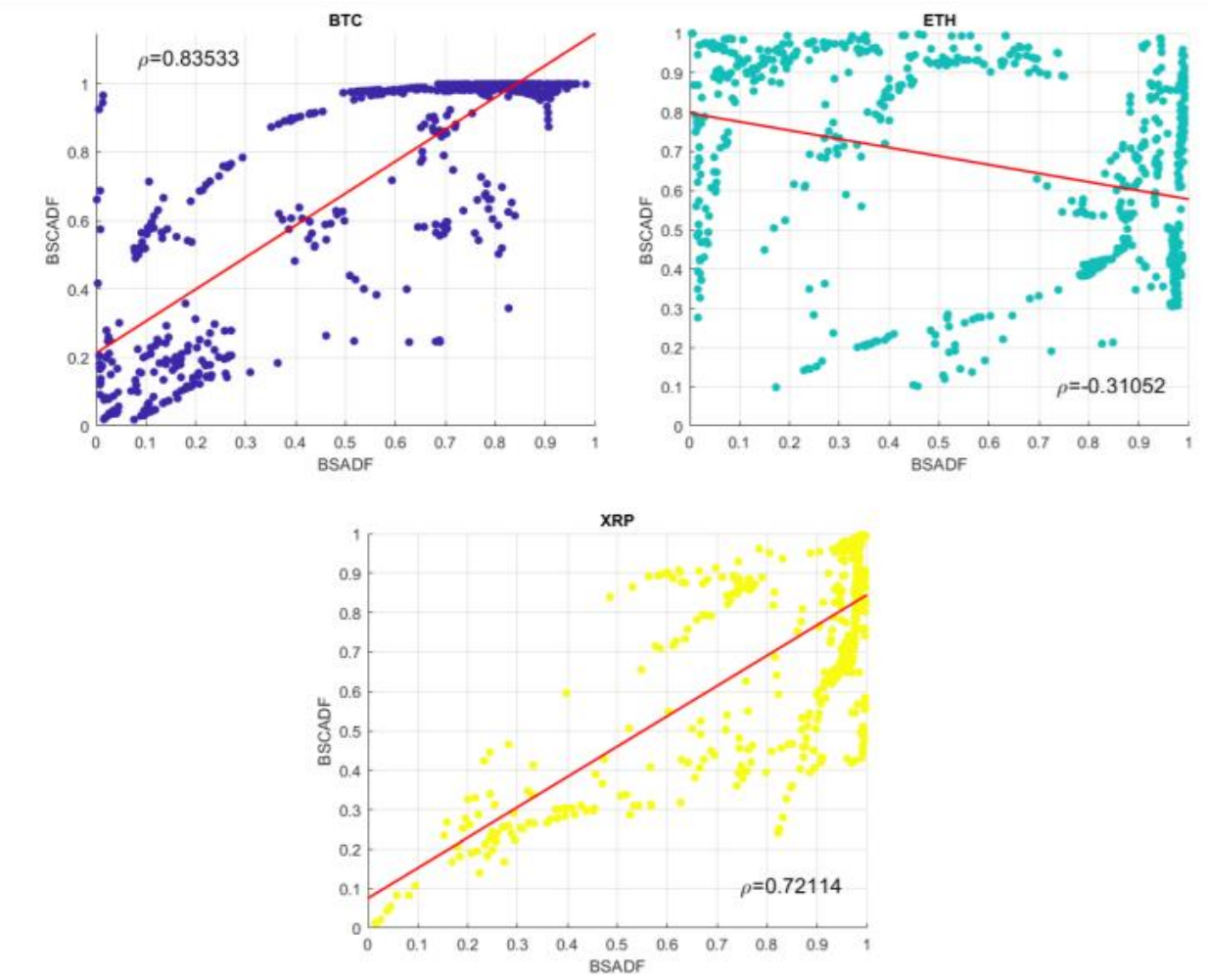


Figure 5: Scatter plot of BSADF and Brain Sentiment Index BSCADF p-values. The figure shows the scatter plot of the BSADF and Brain Sentiment Index BSCADF p-values related for the three cryptocurrencies over the considered sample period. The value of ρ indicates the correlation between the two test statistics.

Conclusions

- **Speculative bubbles** are a key concern to market participants and regulators
- The **occurrence of bubbles** is a frequently investigated phenomenon in the financial and econometric literature
- We propose to exploit the information derived from a large set of cryptocurrency news to **detect** and, possibly, **anticipate** the presence of **speculative bubbles** in cryptocurrency prices
- Through a **Covariate Superior Augmented Dickey-Fuller (CSADF) test**, we exploit market sentiment when testing the presence of explosivity in cryptocurrency prices and we use the test results as an **early warning** indicator
- Our results show that the **Sentiment Indicator** covariate appears to be the one which better foresees the explosive behaviour in cryptocurrency data
- This is arguably due to the fact that volume indicators are, unlike Sentiment indicators, not **polarized**, and are therefore unable to distinguish the possible future trajectory of the price time series