Received: 16 September 2019 | Revised: 15 July 2020 | Accepted: 28 July 2020

DOI: 10.1111/twec.13020

ORIGINAL ARTICLE

The World Economy WILEY

Analysis of Bitcoin prices using market and sentiment variables

Burcu Kapar[1] | **Jose Olmo**[2,3]

[1]American University in Dubai, Dubai, United Arab Emirates

[2]Department of Economic Analysis, Universidad de Zaragoza, Zaragoza, Spain

[3]Department of Economics, University of Southampton, Southampton, UK

KEYWORDS

Bitcoin price, cointegration, cryptocurrency, factor models, permanent-transitory decomposition

1 | INTRODUCTION

There is an emerging empirical stream of literature that has attempted to explain the mechanisms behind cryptocurrency pricing and, in particular, Bitcoin. This paper belongs to this recent tradition; however, in contrast to most empirical studies on the dynamics of the cryptocurrency, we focus on uncovering its efficient price. Informally, the efficient price can be interpreted as an average price that reflects the long-run dynamics of the price after removing short-term fluctuations due to market noise, outliers or the occurrence of other transitory disruptions in the market.

The contribution of this paper is twofold. First, we identify the long-run relationship between Bitcoin and a set of variables with power to explain the dynamics of the cryptocurrency. This is done applying cointegration methods. If Bitcoin is cointegrated with any of the other variables considered, it implies that there is a long-run relationship between them. This can be used by investors to make strategic trading decisions. When there is a deviation from the long-run equilibrium, investors can act on that in the belief that it will return to the long-run equilibrium. The second contribution is to explore the factors with ability to explain the efficient Bitcoin price. More formally, we define the efficient Bitcoin price as the permanent component of the cryptocurrency log-price obtained from the permanent-transitory (P-T) decomposition proposed in Gonzalo and Granger (1995). To the best of our knowledge, this is the first paper that identifies the efficient price of Bitcoin by applying the permanent-transitory (P-T) decomposition proposed by these authors.

Our empirical strategy is based on estimating a vector error correction model (VECM), see Johansen (1988, 1991), to assess the long-run relationship between the Bitcoin spot price and a set of financial, economic and sentiment variables. Motivated by the empirical literature exploring the determinants of Bitcoin and also by the properties of cryptocurrencies, we entertain the following variables: (a) the S&P 500 index capturing the performance of financial markets, (b) the price of gold, a commodity with investment properties similar to those characterising Bitcoin, (c) Google Search as a variable that quantifies the entries in the Google engine associated to Bitcoin, in particular, and cryptocurrencies, in general, and (d) a fear index proxied by the FED Financial Stress Index that captures market sentiment.

The VECM specification is applied to two overlapping periods characterised by very different dynamics. A first period considering data up to January 2018, and a second period that considers data up to May 2019 characterised by a steep decline in the Bitcoin price and a slow recovery. The findings of our empirical study are in stark contrast across periods. For the first period, we find strong positive comovements between the dynamics of Bitcoin price and online interest in the cryptocurrency measured by Google searches. The other factor with a positive relationship with Bitcoin price is the S&P 500 index. Gold spot price and the fear index have, on the other hand, a negative effect. These empirical findings suggest that during this period, Bitcoin is procyclical, driven by investors' interest on the cryptocurrency and positively correlated to the market portfolio. In this period, Bitcoin follows the trend exhibited by financial markets and cannot be considered as an alternative asset or safe haven. These findings are empirically confirmed by the negative relationship between the cryptocurrency, the gold price and the fear index.

These results change dramatically if we include the period 2018–19 in our analysis. Thus, with data spanning the period July 2010 to May 2019, we find that the only variable with power to explain the dynamics of Bitcoin in the long run is online interest in the cryptocurrency. Financial variables such as the S&P 500 index and the gold spot price are no longer statistically significant. The fear index measuring market sentiment is not relevant either. During the last period, Bitcoin prices are mainly driven by individuals' expectations reflected in an increase in online searches followed by a sharp decrease, and not by market fundamentals. In this period, we find statistical evidence of predictability between the returns on Bitcoin and variation in Google searches, that runs in both directions, indicating that these feedback effects act as a mechanism to increase prices based on information (observed interest in the cryptocurrency) and to increase the interest in the cryptocurrency as prices rapidly grow.

We are not the first to apply cointegration and VECM for explaining the dynamics of Bitcoin and interdependencies with other variables. For example, Ciaian, Rajcaniova, and Kancs (2015) propose VECM specifications for Bitcoin prices as a suitable methodology to account for potential endogeneity between the variables, see Lütkepohl and Krätzig (2014). Other recent studies exploiting the cointegration methodology to investigate the long-run relationship between Bitcoin and other cyrptocurrencies are Bação, Duarte, Sebastião, and Redzepagic (2018), Ciaian, Rajcaniova, and Kancs (2018), Leung and Nguyen (2019) and Van den Broek (2018).

The long-term equilibrium relationship between the variables uncovered through the VECM specification is applied to obtain the efficient price of Bitcoin once transitory effects are filtered from observed prices. The Gonzalo–Granger decomposition allows us to differentiate between a permanent component that drives the efficient long-run price of the cryptocurrency and a transitory component. The P-T decomposition also provides an automatic characterisation of the common factors with power to explain the dynamics of the efficient Bitcoin price. These factors are constructed as a linear combination of the set of explanatory variables proposed in the VECM representation and are interpreted in this context as pricing factors. In the empirical exercise corresponding to the period 2010–17, we find four common factors for explaining Bitcoin prices. The first factor can be interpreted as a market portfolio, see Sharpe (1964) and Lintner (1965). The second factor exploits the common features between gold and Bitcoin prices. The third factor is a linear combination of Bitcoin and Google Search, and the fourth factor captures overall market sentiment usually proxied by the FED fear index. These factors track very accurately the permanent component of Bitcoin prices. These factors replace traditional factor models such as Fama and French (1993, 2015) common factors that are shown to perform very poorly in the current setting. These common factors are robust to the sample period; thus, the P-T decomposition obtained from the VECM applied to 2010–19 provides the same set of pricing factors. However, in this case, the contribution of the factors to explaining the efficient Bitcoin price is very different. During this period, the third factor given by a combination of Bitcoin and Google Search gains importance, whereas the market portfolio factor and the factor tracking gold prices lose relevance.

The rest of the paper is organised as follows. Section 2 presents a literature review on the determinants of Bitcoin. The section also describes the data used for our empirical analysis and provides some background on the choice of pricing variables. Section 3 contains the results of the cointegration analysis and derives the equilibrium dynamics of Bitcoin prices after applying a VECM. Section 4 introduces the Gonzalo and Granger (1995) P-T decomposition and derives the efficient Bitcoin price. Section 5 presents a separate empirical asset pricing exercise to illustrate the poor performance of Fama and French type models for explaining Bitcoin. Section 6 concludes the paper with a summary of the main results and proposals for further research. Tables and figures are collected in an Appendix.

2 | LITERATURE REVIEW

The increasing interest in cryptocurrencies has resulted in an exponential growth in the number of studies considering different features of cryptocurrencies. Recent studies differ in the econometric methodologies proposed to study the dynamics of the cryptocurrencies and also on the set of factors used to rationalise its behaviour. We highlight important contributions by Ciaian et al. (2015), Hayes (2017), Kristoufek (2015), Polasik, Piotrowska, Wisniewski, Kotkowski, and Lightfoot (2015), Zhu, Dickinson, and Li (2017), Blau (2017), Zhang, Wang, Li, and Shen (2018), Bouri, Gupta, Lahiani, and Shahbaz (2018), Jareno, Gonzales, Tolentino, and Sierra (2020) and Kalyvas, Papakyriakou, Sakkas, and Urquhart (2020). These authors use a variety of econometric models such as OLS, VECM, Wavelet Analysis, IV Estimation, GARCH and MF-DCCA models and consider market and technical variables to explain the dynamics of cryptocurrencies.

A major feature of the Bitcoin spot market is the importance of unconventional variables such as Google trends and other online and social media activities to explain the dynamics of prices and returns. Several authors such as Kristoufek (2013, 2015), Glaser et al. (2014), Bouoiyour, Selmi, and Tiwari (2014), Gozcek and Skliarov (2019) and Dastgir, Demir, Downing, Gozgor, and Lau (2019) measure individuals' interest on Bitcoin by the increase in the volume of online searches. Similarly, Garcia, Tessone, Mavrodiev, and Perony (2014) show that word-of-mouth information on social media, information on Google Trends and inflows of Bitcoin users have a significant influence on Bitcoin price changes. Bouoiyour and Selmi (2017) use Google search queries for India and Venezuela alongside with some macro and financial factors and find that the growing attention in Bitcoin leads to increasing prices when the market is in normal and bull regimes. By utilising Factor Augmented Vector Error Correction Model, Gozcek and Skliarov (2019) consider large data set of explanatory variables, but their results indicate that the popularity is the main variable that drives Bitcoin prices.

The message emerging from these literatures is that the price of Bitcoin is closely related to individuals' interest in the cryptocurrency and not so much driven by market fundamentals. Dastgir et al. (2019) examine the casual relationship between Bitcoin attention (measured by the Google Trends search queries) and Bitcoin returns for the period from 1 January 2013 to 31 December 2017. By employing the Copula-based Granger Causality in Distribution (CGCD) test, they observe a bi-directional causal relationship between Bitcoin attention and Bitcoin returns especially in the left tail (poor performance) and the right tail (superior performance) of the distribution.

Kristoufek (2015) classifies potential pricing variables into three categories: economic drivers, technical drivers and attractiveness (popularity). Examples of economic drivers measuring the demand for the cryptocurrency are the ratio between trade and exchange transaction volume, the number of transactions excluding exchange transactions, estimated output volume and total bitcoins in circulation. Examples of technical drivers are the hash rate and the difficulty of mining. Variables capturing the popularity of Bitcoin as an investment are the number of Wikipedia searches and metrics reported by Google Trends. Kristoufek (2015)'s study also includes a financial stress index and gold prices to assess the relationship between the cryptocurrency and variables considered as safe havens under financial turmoil. The main conclusion in Kristoufek (2015)'s study is that the price of Bitcoin is mainly driven by investors' interest in the cryptocurrency and not by macroeconomic fundamentals or financial ratios. This observation is most evident in the long run; however, this author also observes a short-term effect characterised by an increasing interest in the cryptocurrency during episodes of explosive prices and during rapid declines that push prices further down. In these scenarios, Bitcoin does not appear to be a safe haven investment.

The most important technical driver of Bitcoin price is the hash rate. This variable captures the mining effort—hashing power—or hash rate, that is measured in gigahashes per second (GH/s). This quantity refers to the computational effort applied by miners to obtain bitcoins over a given time interval, typically 1 day. The hash rate, or number of hashes per second can be thought of as somewhat analogous to the cycles per second (hertz) of computer processors: the higher the hash rate, the more likely it is to successfully mine bitcoins per day, see Hayes (2017). Despite the importance of this factor in the supply of Bitcoins, there is no consensus in the literature about whether it significantly affects the Bitcoin price. Georgoula, Pournarakis, Bilanakos, Sotiropoulos, and Giaglis (2015) and Hayes (2017) find the hash rate to be a significant positive driver of the Bitcoin price. However, Kristoufek (2015), Kjærland, Khazal, Krogstad, Nordstrøm, and Oust (2018) and Bouoiyour and Selmi (2015) find the hash rate to have no statistical significance for modelling Bitcoin price dynamics. The reason for the absence of significance given by these authors is that the hash rate does not affect Bitcoin supply.

In the following subsection, we introduce the data set relevant for our empirical study and motivate the choice of factors with power to explain the dynamics of the cryptocurrency.

2.1 | Data description

Our sample covers weekly data from 22 July 2010 to 19 May 2019. The weekly frequency of the observations is due to availability of data on Google searches on Bitcoin. This time series is reported with this frequency so rather than interpolating in order to construct a time series of daily observations, we prefer to work with all the variables in our empirical specification with weekly frequencies. By doing so, we avoid oversmoothing the daily observations on Google searches, and at the same time, we analyse the market for Bitcoin at a frequency that is informative about market volatility but avoids the noise inherent at high frequencies.

For Bitcoin price data, we use the Coindesk Bitcoin USD Price Index, a simple average of global Bitcoin/USD exchange prices. It is expressed as the midpoint of the bid/ask spread across a number of global exchanges meeting certain criteria with regard to minimum trade size, trading volume and others.[1] Figure A1 below depicts the historical development of the Bitcoin price quoted in USD.

[1]See http://www.coindesk.com/price/bitcoin-price-index for more detailed information. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Bitcoin prices were relatively stable before late 2013. After this date, prices moved quite dramatically reaching values up to $16,858 and reflecting the increased attention of Bitcoin by investors. Also, the market value of all bitcoins in circulation went from about $277,000 in July of 2010 to over $185 billion in January 2018, being at the highest level of $312 billion in December 2017. After January 2018, we observe a steep decrease in the price of Bitcoin and a further recovery during 2019.

Following empirical studies in cryptocurrencies, see also a nice review by Corbet, Lucey, Urquhart, and Yarovaya (2019), we propose a set of macroeconomic and financial variables with the potential to explain the dynamics of Bitcoin. These variables are the S&P 500 index, capturing the performance of global financial markets, the price of gold, as a commodity with investment properties similar to those characterising Bitcoin, Google searches associated to Bitcoin, and a fear index proxied by a financial stress indicator that captures market sentiment. Data on the S&P 500 index, gold prices and the FED Financial Stress Index (fear index) are obtained from the Federal Reserve Bank of St. Louis website.

The influence of stock prices is modelled through the inclusion of the S&P 500 index. Although this financial index is constructed from US equity data, it is widely used as an indicator of the performance of global equity markets and, more generally, global financial markets. We also include gold as a determinant of the dynamics of Bitcoin. This is so because since its inception Bitcoin has reacted positively to negative shocks to stock prices suggesting that Bitcoin is used as a hedging instrument in periods of financial distress. Dyhrberg (2016) tests the hedging capabilities of Bitcoin and shows that Bitcoin has hedging characteristics as gold. Gold also shares other characteristics of cryptocurrencies such as high price volatility and finite supply.

Online activities tracking interest in Bitcoin are also considered by incorporating the weekly time series of Bitcoin views obtained from Google. More precisely, we use the number of times a keyword related to Bitcoin is entered into the Google engine. Google search volume data have been used in many studies as a proxy of users' attractivenes such as Google search and employment (Askitas & Zimmermann, 2009), Google search and consumption expenditures (Vosen & Schmidt, 2011), Google Search and housing prices (Kulkarni, Haynes, Stough, & Paelinck, 2009), Google search and stock trading volume (Vlastakis & Markellos, 2012) and Google Search and Bitcoin prices (Bouoiyour et al., 2014; Kristoufek, 2013, 2015; Fry & Cheah, 2015).

The fourth explanatory variable defining our empirical asset pricing model of Bitcoin prices is a variable measuring overall market sentiment. This is proxied by the FED financial stress index and measures the degree of financial stress in the markets. It is constructed from 18 weekly data series: seven interest rate series, six yield spreads and five other indicators. Each of these variables captures different aspects of stress in financial markets.

We do not include the hash rate as empirical evidence is inconclusive about the statistical contribution of this variable to explain the dynamics of Bitcoin. The following section introduces the econometric specification.

3 | ECONOMETRIC SPECIFICATION

In this section, we introduce the VECM specification for describing the long-run equilibrium relationship between the variables in the pricing model. We do this for two different overlapping periods with very different characteristics. First, we consider data up to January 2018 and, second, we consider data up to May 2019. By doing so, we explore the effect of the turbulent January 2018 to May 2019 period in the modelling of Bitcoin prices within a longer period that contains both calm and turbulent episodes of Bitcoin prices.[2]

The results present contrasting evidence on the relevance of the pricing factors. The VECM is represented as follows:

$$\Delta Y_t = \alpha Z_{t-1} + \sum_{i=1}^{p} \Gamma_i \Delta Y_{t-i} + \epsilon_t, \tag{1}$$

where $Y_t = (p_{\text{Bitcoin}}, p_{\text{S\&P500}}, p_{\text{Gold}}, \text{goo}, \text{fed})'$ with p denoting the log-price of Bitcoin, the S&P 500 index and gold, respectively; goo denotes the log of the variable Google Search and fed is the fear index defined by the FED financial stress indicator. The equilibrium condition is characterised by the error correction variable $Z_t = \beta' Y_t$ with β the cointegrating vector; α is the vector of loadings that measures the speed of adjustment of the departures from equilibrium Z_{t-1} to the dynamics of the vector ΔY_t, with ΔY_t the variables in first differences. The matrices Γ_i capture the short-run Granger-casual relationships between the variables in first differences. The random vector ϵ_t denotes the multivariate error term.

3.1 | Period 2010–17

Summary statistics of all the variables considered in our study during this period are reported in the top panel of Table A1. The maximum price of Bitcoin in our sample is \$16,858 and the minimum is \$0.05. Bitcoin has by far the highest level of volatility across all series. It is positively skewed and exhibits an excess kurtosis with respect to the normal distribution. These empirical features of the series of Bitcoin prices are quite different from the rest of variables in the system. For example, the S&P 500 index and the price of gold have a small skewness parameter and the kurtosis is close to that of the standard normal distribution. Only the variable Google Search has summary statistics similar to those of the time series of Bitcoin prices.

Table A2 presents the results of unit root and cointegration tests for the period July 2010 to January 2018 using several vector autoregressive specifications up to five lags. The left panel presents the empirical results for the Augmented Dickey–Fuller (ADF) and Phillips and Perron (P-P) tests. Both tests fail to reject the null hypothesis of non-stationarity, and hence, each series contain unit root. The middle panel presents the p-value of the Engle–Granger two-stage cointegration test. A low p-value of the test entails the rejection of the null hypothesis and provides evidence of cointegration between the series. The right panel reports the p-value and log-likelihood ratio test for the Johansen's test introduced in Johansen (1988, 1991). The results provide evidence of a single cointegration relationship between the variables in the system.

We retain the VECM with one lag and present the subsequent analyses using this model. The cointegrating vector β characterises the following long-term relationship

$$Z_t = p_{\text{Bitcoin}} - 0.98)3.48 p_{S\&P500} + 1.08)3.22 p_{\text{Gold}} - 0.06)0.71 \text{goo} + 0.35)0.80 \text{fed}. \tag{2}$$

[2]The choice of these two periods is not determined by a formal statistical analysis derived from the application of structural break tests. We indeed carry out these tests over the entire sample period and find a structural break at January 2018, among a few other periods. However, rather than splitting the sample into two very unequal disjoint periods, we prefer to study the effect of the period January 2018 to May 2019 within a longer period that contains both calm and turbulent episodes of Bitcoin prices.

Operating with this expression, we obtain the following equation that describes the price of Bitcoin in equilibrium:

$$p_{\text{Bitcoin}} = 0.98)3.48 p_{\text{S\&P500}} - 1.08)3.22 p_{\text{Gold}} + 0.06)0.71 \text{goo} - 0.35)0.80 \text{fed}. \tag{3}$$

An increase in the S&P 500 index or in Google search queries is corresponded by an increase in Bitcoin prices. In contrast, an increase in gold prices or the fear index is corresponded by a fall in the Bitcoin price.

An interesting implication of the vector error correction model is the possibility of testing for long-run causality from the equilibrium condition Z_{t-1} to the series of Bitcoin prices. More formally, a statistically significant parameter α suggests that departures of Z_t from zero have predictive ability on the dynamics of the return process ΔY_t. That is, a negative value of the parameter α entails mean reversion to the equilibrium condition (3). The magnitude of the parameter α conveys information about the speed of adjustment of the time series ΔY_t to restore equilibrium in the system. The vector of α coefficients corresponding to the above VECM with one lag is

$$\alpha = ((0.016))-0.041, (0.0018))0.0045, (0.002))-0.004, (0.034))0.100, (0.463))-0.004)'. \tag{4}$$

These results suggest that departures from equilibrium, given by values of Z_t different from zero, have in general a mean-reverting effect on the dynamics of the different series. A positive value of Z_t can be interpreted as evidence of a positive idiosyncratic shock in the cryptocurrency market. According to the VECM (1) and the parameter estimates (4), this positive value results in a negative correction of the price of Bitcoin ($\alpha_1 = -0.041$) and Gold ($\alpha_3 = -0.004$) in the next period, and an increase in the price of the S&P 500 index ($\alpha_2 = 0.0045$) and online interest in the cryptocurrency ($\alpha_4 = 0.100$). In contrast, the fear index can be considered exogenous to departures from the equilibrium condition (3) since the parameter α_5 is not statistically significant.[3]

Table A3 reports the p-values of the Lagrange Multiplier test for autocorrelation of the error terms. Overall, the results are very supportive of the VECM specification with one lag. For completeness, Table A5 reports the p-values of the test for the presence of heteroscedasticity in the residuals of the fitted models. In this case, the results suggest that the proposed econometric specification could be further improved by including a model for conditional heteroscedasticity of the error terms. Accounting properly for conditional heteroscedasticity in the error terms is important for understanding the dynamics of the volatility process of Bitcoin over time. Econometrically, it is important for the efficient estimation of the model parameters; however, it does not affect consistency of the parameter estimates. In our setting, we are interested in developing a pricing model for Bitcoin similar in spirit to models proposed in the empirical finance literature using observed and unobserved factors for pricing risky assets.

3.2 | Period 2010–19

This block is divided into two different analyses of cointegration. First, we repeat the above analysis for the extended sample period covering data up to May 2019; second, we also carry out an analysis of cointegration for this period only considering Google searches and Bitcoin prices.

[3]Unreported results available from the authors upon request show that only gold prices Granger-cause the Bitcoin price. This concept of causality refers to short-term causality, that is to the existence of predictability from the lagged returns on gold on the series of Bitcoin financial returns.

Summary statistics of all the variables considered in our study during this period are reported in the bottom panel of Table A1. The Dickey–Fuller test fails to reject the null hypothesis of unit root for any of the variables. The full analysis of cointegration between the five variables discussed above is repeated now for an extended period that includes turbulence in the cryptocurrency market. Interestingly, for this period, the long-run equilibrium condition becomes

$$Z_t = p_{\text{Bitcoin}} + 0.64)0.96 p_{\text{S\&}P500} + 0.45)2.06 p_{\text{Gold}} - 0.00)1.97 \text{goo} + 0.21)1.32 \text{fed}. \tag{5}$$

The only variable that is statistically related in the long run to Bitcoin prices is Google Search. This result illustrates analytically the decoupling between Bitcoin and economic and sentiment variables during this period. The vector with the speed of adjustment parameters is $\alpha = (0.006)0.004, 0.483)0.000, 0.125)-0.000, 0.64)0.96, 0.000)0.031, 0.083)-0.002)$. These results clearly reveal the exogenous character of the variables in the cointegration equation. The only endogenous variables in the VECM specification are the time series of Bitcoin prices and Google Search. This scenario agrees with the empirical literature mentioned above that stresses the role of online searches as the only variable that can rationalise the long-run dynamics of Bitcoin price.

In order to study in more detail the statistical ability of online searches in explaining Bitcoin prices, we carry out a separate analysis of cointegration between the two variables. Figure A1 plots the dynamics of both variables over the period under study. The plots reveal strong positive comovements between both series. Interestingly, although non-stationary, online searches tend to revert more quickly to the mean than Bitcoin prices. This observation may be driven by a lower volatility of the time series of Google searches than Bitcoin prices. Economically, it suggests that a positive outlook on Bitcoin favours investors' interest in the cryptocurrency that is reflected in increasing prices. This online interest on the cryptocurrency recedes over time; however, Bitcoin prices take more time to react and decrease. This is observed during the years 2014 and 2015 reported in Figure A1 and, more recently, in the significant drop in Bitcoin prices after 2018.

We formally assess the relationship between these variables by fitting a VECM. We find that a VECM (1) is sufficient to capture the long-run relationship between the variables. The equilibrium condition between both variables implies the relationship

$$p_{\text{Bitcoin}} = 3.98 + 0.000)1.93 \text{goo}.$$

Interestingly, the speed of adjustment parameter α is in this case $\alpha = (0.519)-0.005, 0.000)0.055)$, indicating that Google Search adjusts to departures from equilibrium between the two variables but Bitcoin price is exogenous. For example, if Bitcoin price goes up beyond what the equilibrium relationship dictates, the model predicts a further increase in Google searches in the next period that is driven by the parameter 0.055. The same result is observed for negative shocks to Bitcoin price affecting the equilibrium condition.

The analysis of short-run predictability reveals the presence of Granger causality in both directions. An increase in online searches has a positive and significant effect on next period's log return on Bitcoin—the parameter estimate is 0.094 with a p-value of .003—and an increase in the log return on Bitcoin has a positive impact in the next period on the volume of Google searches—the parameter estimate is 0.516 with p-value equal to zero.[4] These feedback effects act as a mechanism to increase prices based on individuals' information gathered through online searches, and to increase the interest in the cryptocurrency as individuals observe Bitcoin prices rapidly grow.

[4]Full VECM parameter estimates for this model are available from the authors upon request.

4 | EMPIRICAL PRICING MODEL

This section introduces the Gonzalo–Granger P-T decomposition as a mechanism to obtain endogenously the set of pricing factors with power to explain the efficient Bitcoin price. Following the literature on price discovery initiated by Garbade and Silber (1983), we interpret the efficient price as the permanent component of Bitcoin price. As in previous sections, we divide the analysis into two overlapping periods.

4.1 | Gonzalo–Granger P-T decomposition

In the Gonzalo and Granger (1995) decomposition, the permanent component is characterised by a set of common factors f_t that exhibit unit root behaviour and drive the long-term dynamics of Bitcoin prices. Gonzalo and Granger (1995) impose the following two conditions to identify the common factors: (a) the factors are linear combinations of the original variables Y_t and (b) the transitory component does not have predictive ability on the common factors. Under these two conditions the only linear combinations of the original variables such that the transitory component has no long-run impact on them are

$$f_t = \alpha'_{\perp} Y_t, \tag{6}$$

with $\alpha_{\perp}$defined in our example by a 5×4 matrix satisfying $\alpha'_{\perp}\alpha = 0$. Once the common factors $f_t$are identified the P-T decomposition of $Y_t$is obtained as a linear combination of the factors $f_t$with loadings $A_1$and a transitory component given by the cointegrating vector $Z_t$with loadings A_2. More formally, for the vector $Y_t$the corresponding P-T decomposition is

$$Y_t = A_1 f_t + A_2 Z_t, \tag{7}$$

with $A_1 = \beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}$ and $A_2 = \alpha(\beta'\alpha)^{-1}$, where $\beta_{\perp}$is a 5×4 matrix satisfying that $\beta'_{\perp}\beta = 0$. The factors $f_t$contain the linear combinations of ΔY_tthat have the common feature of not containing the levels of the error correction term $Z_t$in them. Let $Y_t^* = A_1 f_t$denote the unobserved permanent component associated to the original vector $Y_t$that we interpret as the efficient price—also called fundamental value—of Bitcoin.

4.2 | Period 2010–17

The left panel of Table A5 reports the coefficients of the matrix $\alpha_{\perp}$ defining the common pricing factors during this period. The right panel reports the coefficients of the factor loadings A_1 characterising the permanent component of Y_t. The reported parameter values are constructed using the estimates of the vector error correction model specification fitted in the previous section. The results show that there are four common factors driving the dynamics of the five time series. The first column on the left panel of Table A5 shows that the first common factor is mainly driven by the dynamics of the S&P 500 index. The second column reports the parameters that define the second common factor. This factor is mainly driven by the dynamics of gold prices. The third column shows that the third common factor is a combination of Bitcoin and Google Search. The fourth factor is mainly driven by the value 0.998 that is associated to the fear index with the remaining variables having a negligible contribution on the factor structure.

The first factor can be interpreted as the market portfolio as in standard asset pricing formulations, see Sharpe (1964), Lintner (1965) and, more recently, Fama and French (1993, 2015). In contrast to these formulations, our market portfolio factor is driven by the price of the S&P 500 index and not by the excess returns. The second factor exploits the common features between gold and Bitcoin prices. The third factor constitutes one of the main distinctive features of our empirical analysis. This factor is a linear combination of Bitcoin and Google Search. Visual inspection of the joint dynamics of these variables in Figure A1 shows evidence of a strong positive comovement between both variables. The price of Bitcoin responds very quickly to the attention of investors to the cryptocurrency such that it increases when the volume of Google searches grows, and rapidly decreases when interest in the cryptocurrency drops. This factor is the main instrument for explaining the excess kurtosis observed in Bitcoin prices. The other factors are unable to replicate Bitcoin large fluctuations in returns. The fourth factor captures overall market sentiment as proxied by the FED fear index.

The analysis of the matrix A_1 on the right panel of Table A5 provides information on the specific composition of the variables Y_t^*. In particular, we are interested in the first component Y_{1t}^* that represents the efficient Bitcoin price. This variable is defined by the vector of parameters $(1.009, -0.934, 0.681, -0.250)$ such that

$$Y_{1t}^* \equiv p_{\text{Bitcoin}}^* = 1.009 f_{1t} - 0.934 f_{2t} + 0.681 f_{3t} - 0.250 f_{4t}. \tag{8}$$

The magnitude of the parameters associated to the four pricing factors reveal an important effect on the efficient Bitcoin price. Combining this information with the estimates of $\alpha_\perp$ on the left panel, we conclude that the efficient price is positively related to the S&P 500 index and negatively related to gold prices. The fear index also has a negative effect on the long-run dynamics of Bitcoin. The top panel of Figure A2 reports the dynamics of the efficient price $Y_{1t}^* \equiv p_{\text{Bitcoin}}^*$ obtained from the P-T decomposition along with the dynamics of the original Bitcoin price denoted by Y_{1t}. The bottom panel reports the permanent component of the variable Google Search Y_{4t}^* along with the original time series Y_{4t}.

The other series are not reported to save space but are available from the authors upon request. In all cases the permanent components—efficient price in the case of Bitcoin—are very close to the actual time series. A more detailed analysis of the P-T decomposition also allows us to analyse the factors that drive the permanent component of the remaining variables in the system, namely S&P 500 index, gold and the fear index. The parameters in the second row of A_1 describe the factor loadings defining the permanent component of the S&P 500 index. The relevant parameters are $(0.893, 0.098, -0.001, 0.023)$ that suggest that the fundamental value of this financial index is mainly driven by the first common factor. Note, however, that this factor was constructed from the S&P 500 index. These two facts suggest that the S&P 500 index is exogeneous and, hence, the common factor tracking this variable is, in fact, a close representation of the variable itself. Similar results are obtained for the fundamental values of gold and the fear index. The analysis of Google Search is different. In this case, the permanent component tracking this variable is driven by a combination of the four pricing factors, with loadings given by $(-2.360, 2.183, 0.964, 0.529)$.

4.3 | Period 2010–19

Table A6 reports the coefficients of the matrix $\alpha_\perp$and the factor loadings A_1characterising Y_t^*through the expression $Y_t^* = A_1 f_t$, with $f_t = \alpha_\perp' Y_t$. The reported parameter values are similar to those obtained for the previous decomposition. There are small differences in the construction of the common factors but the overall definition of the pricing factors is robust across sample periods. The first factor can be interpreted as a market portfolio. The second factor is mainly driven by gold prices. The third factor is a linear combination of Bitcoin and Google Search, and the fourth factor is driven by the FED fear index. The construction of the efficient price of Bitcoin $Y^*_{1t}$is, however, very different from the previous case. The inclusion of the period 2018–19 modifies completely the sensitivity of the efficient Bitcoin price to the common factors. The permanent component of Bitcoin is now defined as

$$Y^*_{1t} \equiv p^*_{\text{Bitcoin}} = -0.153f_{1t} - 0.329f_{2t} + 0.879f_{3t} - 0.249f_{4t}. \tag{9}$$

The sensitivity of the first factor swaps sign with respect to the analysis of 2010–17 and loses relevance for describing the efficient Bitcoin price. The magnitude of the second factor is also significantly smaller. The third factor, driven by a linear combination of the Bitcoin time series and Google Search, gains importance in this period. This result is consistent with the findings of the reduced VECM (1). The contribution of the fear index factor is, however, similar to the previous period.

Figure A3 illustrates the performance of the P-T decomposition in approximating the time series of Bitcoin prices and Google Search. In particular, the top panel of Figure A3 reports the dynamics of Y^*_{1t}, tracking the efficient Bitcoin price, and the bottom panel reports the dynamics of Y^*_{4t}, tracking the permanent component of Google Search. The remaining examples are available from the authors upon request.

5 | AN ASSET PRICING MODEL FOR BITCOIN RETURNS

The empirical analysis carried out above presents two equations for explaining the price of Bitcoin. Equation (3) describes the actual price of Bitcoin as a linear combination of the S&P 500 index, gold, Google Search and the FED financial stress indicator. This pricing equation is obtained under equilibrium conditions using a VECM representation of the relationship between the five variables in the system. The VECM also provides an equation for explaining the short-run dynamics of the returns on Bitcoin around the equilibrium condition. In particular, rewriting Equation (1), we have

$$r_{t,\text{Bit}} = -0.016)0.041Z_{t-1} + 0.051)0.15r_{t-1,\text{Bit}} - 0.334)0.59r_{t-1,\text{gold}} + \epsilon_{t,\text{Bit}}, \tag{10}$$

with $\epsilon_{t,\text{Bit}}$ the error term associated to the Bitcoin time series. The return on Bitcoin is mean-reverting such that positive departures from the equilibrium condition, interpreted as abnormal movements in the Bitcoin price, entail a negative return on the cryptocurrency in the next period. There is also evidence of serial persistence of Bitcoin returns and negative predictive ability of the return on gold on next period Bitcoin returns.

Alternatively, we can also derive a pricing equation taking first differences in Equation (8). In this case, we define the efficient Bitcoin return, that is defined as

$$r^*_{\text{Bit}} = 1.009r_{f_1,t} - 0.934r_{f_2,t} + 0.681r_{f_3,t} - 0.250r_{f_4,t}, \tag{11}$$

where $r_{f_i,t}$ for $i = 1, \cdots, 4$ denote the first differences on the pricing factors obtained from the P-T decomposition. This expression can be interpreted as a pricing equation similar in spirit to standard empirical asset pricing models such as Fama and French (1993, 2015) and Carhart (1997). As noted earlier, the return on the first factor, $r_{f_1,t}$, is very close to the return on the S&P 500 index. This fact allows us to interpret the pricing formula (11) as an extension of the standard CAPM that incorporates a different set of pricing factors constructed endogenously from the P-T decomposition. The second factor is characterised by the return on gold, the third factor is a combination of the growth in Google searches and Bitcoin returns, and the fourth factor denotes the growth of the fear index.

5.1 | Fama-French specification

The empirical application is completed by presenting an empirical asset pricing exercise. In this case, the interest is not in discovering the Bitcoin price but, instead, in testing statistically the predictive power of standard pricing factors. In particular, we focus on the Fama-French three-factor model that is fitted in Table A7 to the three time series of returns introduced above.

First, we consider the excess return on Bitcoin given by $r_{\mathrm{Bit}} - r_{0t}$, with r_{0t} the weekly risk-free interest rate proxied by the 3-month US Treasury bill. Second, we consider the fundamental value of the excess return on Bitcoin given by $r^*_{\mathrm{Bit}} - r_{0t}$. Third, we consider the excess returns on the transitory component of Bitcoin. The latter component is stationary by construction and can be considered as a return on the asset. The transitory component is estimated as the difference between p_{Bit} andp^*_{Bit}.[5] The results show the poor performance of the Fama-French three-factor model for explaining the dynamics of the excess returns on Bitcoin, $r_{\mathrm{Bit}} - r_{0t}$, for the time series under study. None of the pricing factors are statistically significant and the coefficient of determination R^2 is very close to zero.

The results for the three-factor model for the excess returns $r^*_{\mathrm{Bit}} - r_{0t}$ yield similar results. In this case, the coefficient associated to the excess return on the market portfolio is positive and statistically significant highlighting the role of this pricing factor for explaining the fundamental value of Bitcoin spot prices. This finding is very much in the spirit of Sharpe's CAPM. In our asset pricing framework, we rationalise this result as a consequence of the long-run relationship between the financial index and Bitcoin prices shown above. For completeness, we have also fitted the three-factor model to the transitory component of Bitcoin prices. The results show the poor predictive ability of the Fama-French pricing factors for explaining the transitory component of Bitcoin.

6 | CONCLUSIONS

This paper provides an empirical exercise of Bitcoin price discovery. To do this, we propose a VECM and a permanent-transitory decomposition of the Bitcoin spot price series. This modelling strategy provides a set of common factors for describing the efficient price—permanent component—of the cryptocurrency that is endogenously constructed as a combination of a set of explanatory variables.

The results of the model applied to data from July 2010 to January 2018 show the relevance of four pricing factors for explaining the efficient Bitcoin price. These factors are the S&P 500 index, the price of gold, Google online searches associated to Bitcoin and a fear index proxied by the FED Financial Stress Index that captures market sentiment. The S&P 500 index has a positive effect on the fundamental value of Bitcoin prices and the gold spot price and the fear index have a negative effect. These empirical findings suggest that during this period Bitcoin is procyclical, driven by investors' interest on the cryptocurrency and positively correlated to the market portfolio. In this period, Bitcoin follows the trend exhibited by financial markets and cannot be considered as an alternative asset or safe haven.

[5]Alternatively, the transitory component can be obtained, see (7), as $\alpha_1(\beta'\alpha)^{-1}Z_t$ with α_1 the first element of the vector α carrying the long-term causality in (1).

We also fit the model to data from July 2010 to May 2019. The latter period is characterised by a steep drop in price. Our econometric specification shows that the only variable with power to explain the long-run dynamics of the cryptocurrency in this period is Google Search, a variable that measures online interest in the cryptocurrency. The economic and sentiment variables are not statistically significant. The permanent-transitory decomposition shows the robustness of the common factors for explaining Bitcoin's efficient price during this period. The main difference with respect to the previous period is that the only factor that really contributes to explaining the efficient cryptocurrency price is the factor that combines the Bitcoin price and Google Search. This variable can be considered as a very good predictor of the future dynamics of the cryptocurrency.

Future research on the topic plans to extend the current empirical model by investigating some other tools such as Wikipedia search, newspaper news etc for explaining Bitcoin price dynamics.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

ORCID

Jose Olmo https://orcid.org/0000-0002-0437-7812

REFERENCES

Askitas, N., & Zimmermann, K. F. (2009). Google econometrics and unemployment forecasting. *Applied Economics*, *55*(2), 107–120. https://doi.org/10.3790/aeq.55.2.107

Bação, P., Duarte, A. P., Sebastião, H., & Redzepagic, S. (2018). *Information transmission between cryptocurrencies: Does bitcoin rule the cryptocurrency world?* Working Paper of Centre for Business and Economics Research. https://ideas.repec.org/p/gmf/papers/2018-06.html

Blau, B. (2017). Price dynamics and speculative trading in bitcoin. *Research in International Business and Finance*, *41*, 493–499. https://doi.org/10.1016/j.ribaf.2017.05.010

Bouoiyour, J., & Selmi, R. (2015). What does bitcoin look like? *Annals of Economics and Finance*, *16*(2), 449–492.

Bouoiyour, J., & Selmi, R. (2017). *The bitcoin price formation: Beyond the fundamental sources*. Working Paper of Cornell University. https://arxiv.org/ftp/arxiv/papers/1707/1707.01284.pdf

Bouoiyour, J., Selmi, R., & Tiwari, A. (2014). *Is bitcoin a business income or speculative bubble? Unconditional vs. conditional frequency domain analysis*. Working Paper of Munich Personal RePEc Archieve. https://mpra.ub.uni-muenchen.de/59595/

Bouri, E., Gupta, R., Lahiani, A., & Shahbaz, M. (2018). Testing for asymmetric nonlinear short- and long-run relationships between bitcoin, aggregate commodity and gold prices. *Resources Policy*, *57*, 224–235. https://doi.org/10.1016/j.resourpol.2018.03.008

Carhart, M. M. (1997). On persistence of mutual fund performance. *The Journal of Finance*, *52*, 57–82.

Ciaian, P., Rajcaniova, M., & Kancs, D. (2015). The economics of bitcoin price formation. *Appplied Economics*, *48*(19), 1799–1815. https://doi.org/10.1080/00036846.2015.1109038

Ciaian, P., Rajcaniova, M., & Kancs, D. (2018). Virtual relationships: Short and long-run evidence from bitcoin and altcoin markets. *Journal of International Financial Markets Institutions and Money*, *52*, 173–195. https://doi.org/10.1016/j.intfin.2017.11.001

Corbet, S., Lucey, B., Urquhart, A., & Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, *62*, 182–199. https://doi.org/10.1016/j.irfa.2018.09.003

Dastgir, S., Demir, E., Downing, G., Gozgor, G., & Lau, C. K. M. (2019). The causal relationship between bitcoin attention and bitcoin returns: Evidence from the Copula-based Granger Causality Test. *Finance Research Letters*, *28*, 160–164. https://doi.org/10.1016/j.frl.2018.04.019

Dyhrberg, A. H. (2016). Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters*, *16*, 139–144. https://doi.org/10.1016/j.frl.2015.10.025

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*, 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*, 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fry, J., & Cheah, E. (2015). Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Garbade, K. D., & Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, *65*(2), 289–297. https://doi.org/10.2307/1924495

Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: Feedback cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface*, *11*, 1–8. https://doi.org/10.1098/rsif.2014.0623

Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D. N., & Giaglis, G. M. (2015). *Using time-series and sentiment analysis to detect the determinants of bitcoin prices*. Retrieved from https://ssrn.com/abstract=2607167

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M., Moritz, C., & Siering, M. (2014). *Bitcoin – Asset or currency? Revealing users' hidden intentions*. Retrieved from https://ssrn.com/abstract=2425247

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business and Economic Statistics*, *13*, 27–35. https://doi.org/10.1080/07350015.1995.10524576

Gozcek, L., & Skliarov, I. (2019). What drives the bitcoin price? A factor augmented error correction mechanism investigation. *Applied Economics*, *51*, 6393–6410. https://doi.org/10.1080/00036846.2019.1619021

Hayes, A. (2017). Cryptocurrency value formation: An empirical analysis leading to a cost of production model for valuing bitcoin. *Telematics and Informatics*, *34*(7), 1308–1321. https://doi.org/10.2139/ssrn.2648366

Jareno, M., Gonzales, M., Tolentino, M., & Sierra, K. (2020). Bitcoin and gold price returns: A quantile regression and NARDL analysis. *Resources Policy*, *67*, 101666. https://doi.org/10.1016/j.resourpol.2020.101666

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, *59*, 1551–1580. https://doi.org/10.2307/2938278

Kalyvas, A., Papakyriakou, P., Sakkas, A., & Urquhart, A. (2020). What drives bitcoin's price crash rRisk? *Economics Letters*, *191*, 108777.

Kjærland, F., Khazal, A., Krogstad, E. A., Nordstrøm, F. B. G., & Oust, A. (2018). An analysis of bitcoin's price dynamics. *Journal of Risk and Financial Management*, *11*(4), 63. https://doi.org/10.3390/jrfm11040063

Kristoufek, L. (2013). BitCoin meets Google trends and Wikipedia: Quantifying the relationship between phenomena of the internet era. *Scientific Reports*, *3*, 3415. https://doi.org/10.1038/srep03415

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

Kulkarni, R., Haynes, K., Stough, R., & Paelinck, J. H. P. (2009). *Forecasting housing prices with Google econometrics* (GMU School of Public Policy Research Paper No. 2009-10).

Leung, T., & Nguyen, H. (2019). Constructing cointegrated cryptocurrency portfolios for statistical arbitrage. *Studies in Economics and Finance*, *36*(3), 581–599. https://doi.org/10.1108/SEF-08-2018-0264

Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in the portfolios and capital budgets. *The Review of Economics and Statistics*, *47*, 13–37.

Lütkepohl, H., & Krätzig, M. (2014). *Applied time series econometrics*. Cambridge, UK: Cambridge University Press.

Polasik, M., Piotrowska, A., Wisniewski, T., Kotkowski, R., & Lightfoot, G. (2015). Price fluctuations and the use of bitcoin: An empirical inquiry. *International Journal of Electronic Commerce*, *20*(1), 9–49.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*, 425–462. https://doi.org/10.2307/2977928

Van den Broek, L. (2018). *Cointegration-based pairs trading framework with application to the cryptocurrency market*. Bachelor Thesis, Erasmus University Rotterdam.

Vlastakis, N., & Markellos, R. (2012). Information demand and stock market volatility. *Journal of Banking and Finance*, *36*(6), 1808–1821. https://doi.org/10.1016/j.jbankfin.2012.02.007

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google trends. *Journal of Forecasting*, *30*(6), 565–578. https://doi.org/10.1002/for.1213

Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Quantifying the cross-correlations between online searches and bitcoin market. *Physica A: Statistical Mechanics and Its Applications*, *509*, 657–672. https://doi.org/10.1016/j.physa.2018.06.073

Zhu, Y., Dickinson, D., & Li, J. (2017). Analysis on the influence factors of bitcoin's price based on VEC model. *Financial Innovation*, *3*(1), 1–13.

How to cite this article: Kapar B, Olmo J. Analysis of Bitcoin prices using market and sentiment variables. *World Econ*. 2021;44:45–63. https://doi.org/10.1111/twec.13020

APPENDIX

TABLE A1 Summary statistics for the variables in the study

	Variables	N	Mean	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
2010–18	Bitcoin price	392	845	2,252	0.05	16,858	4.98	30.2
	S&P 500 index	392	1,799	427	1,047	2,798	0.07	1.93
	Gold price	392	1,364	191.3	1,049	1,855	0.75	2.43
	Google search	392	4.58	10.93	0	100	5.80	42.02
	Fear index	392	−1.06	0.33	−1.58	−0.081	0.54	2.51
2010–19	Bitcoin price	461	1,685	3,003	0.05	19,086	2.39	9.12
	S&P 500 index	461	1,942	523	1,048	2,943	0.15	1.92
	Gold price	461	1,352	181	1,056	1,897	0.98	2.96
	Google search	461	5.82	10.8	0	100	5.02	35.2
	Fear index	461	−1.00	0.33	−1.53	0.001	0.64	2.76

TABLE A2 p-Values are reported corresponding to different unit root and cointegration tests for the time series of prices over the period from July 2010 to January 2018

	ADF and Philips Perron unit root test						Engle-Granger(H_0: no coint)	Johansen test (H_0: rank = r)	
Lags		Bitcoin	S&P 500	Gold	Google	Fear		Trace statistics	Log-lik.
1	ADF	0.40	0.25	0.15	0.13	0.20	0.03	0 87.8	2,614.9
1	P-P	0.30	0.18	0.23	0.14	0.18		1 48.8	2,634.4
2	ADF	0.21	0.20	0.22	0.12	0.23	0.04	0 74.4	2,672.3
2	P-P	0.28	0.16	0.24	0.14	0.17		1 43.0	2,688.0
3	ADF	0.21	0.21	0.24	0.11	0.27	0.08	0 64.6	2,687.7
3	P-P	0.28	0.17	0.24	0.14	0.17		1 38.0	2,701.0
4	ADF	0.17	0.26	0.31	0.10	0.22	0.12	0 61.0	2,694.9
4	P-P	0.26	0.18	0.24	0.14	0.17		1 36.7	2,707.0
5	ADF	0.10	0.23	0.25	0.09	0.21	0.09	0 61.9	2,696.9
5	P-P	0.25	0.19	0.24	0.14	0.17		1 37.7	2,709.0

TABLE A3 Lagrange Multiplier test for autocorrelation of the errors vector error correction representation (1) with one lag for the sample covering the period July 2010 to January 2018

Lag	Chi^2	*df*	Prob > chi^2
1	33.7	25	0.11
2	31.4	25	0.17
3	22.7	25	0.59
4	30.5	25	0.20
5	23.7	25	0.53

TABLE A4 Residual heteroskedasticity test for the vector error correction representation (1) with one lag for the sample covering the period July 2010 to January 2018

	Chi-sq	*df*	Prob.
No cross term	547.6	180	0.00
With cross term	1,003.7	405	0.00

TABLE A5 Parameter estimates corresponding to the Gonzalo-Granger Permanent-Transitory Decomposition for the analysis of Bitcoin. The sample period comprises weekly data from July 2010 to January 2018. Left panel reports the parameters of the matrix $\alpha_\perp$ in (6) and the right panel reports the parameters of the matrix A_1 in (7)

$\alpha_\perp$				A_1 loadings			
0.041	−0.038	0.921	−0.042	1.009	−0.934	0.681	−0.250
0.998	0.001	−0.027	0.001	0.893	0.098	−0.001	0.023
0.001	0.998	0.025	−0.001	0.099	0.907	0.001	−0.022
−0.027	0.025	0.385	0.028	−2.360	2.183	0.964	0.529
0.001	−0.001	0.028	0.998	0.109	−0.101	0.001	0.975

TABLE A6 Parameter estimates corresponding to the Gonzalo-Granger Permanent-Transitory Decomposition for the analysis of Bitcoin. The sample period comprises weekly data from July 2010 to May 2019. Left panel reports the parameters of the matrix $\alpha_\perp$ in (6) and the right panel reports the parameters of the matrix A_1 in (7)

$\alpha_\perp$				A_1 loadings			
−0.013	−0.28	0.929	−0.078	−0.153	−0.329	0.879	−0.249
0.999	−0.0002	0.008	−0.0007	0.994	−0.011	0.007	−0.007
−0.0002	0.999	0.019	−0.001	−0.011	0.975	0.015	−0.015
0.008	0.019	0.363	0.053	0.373	0.800	0.493	0.497
−0.0007	−0.001	0.053	0.995	−0.031	−0.067	0.042	0.957

TABLE A7 Fama-French three-factor models with monthly data covering the period July 2010 to January 2018

Variables	$r_{\text{Bit}} - r_0$	$r^*_{\text{Bit}} - r_0$	$(p_{\text{Bit}} - p^*_{\text{Bit}}) - r_0$
Mkt-r_0	0.434	0.635	−0.259
	(0.347)	(0.105)	(0.533)
SMB	−0.266	0.337	−0.584
	(0.698)	(0.642)	(0.372)
HML	−0.871	0.068	0.996
	(0.193)	(0.903)	(0.129)
Constant	2.749***	2.543***	21.448***
	(0.000)	(0.001)	(0.000)
Observations	389	389	389
R-squared	.005	.009	.011

Note: r_{Bit} denotes the return on Bitcoin prices, r^*_{Bit} denotes the return on the fundamental value of Bitcoin and $p_{\text{Bit}} - p^*_{\text{Bit}}$ denotes the transitory component of Bitcoin prices. Mkt denotes the market portfolio and r_0 is the risk-free rate. SMB stands for the small-minus-big portfolio and HML stands for the high-minus-low portfolio. Robust *p*-values are in parentheses.

*** denotes *p-value* < .01, ** denotes *p-value* < .05, and **denotes p-value* < .1.

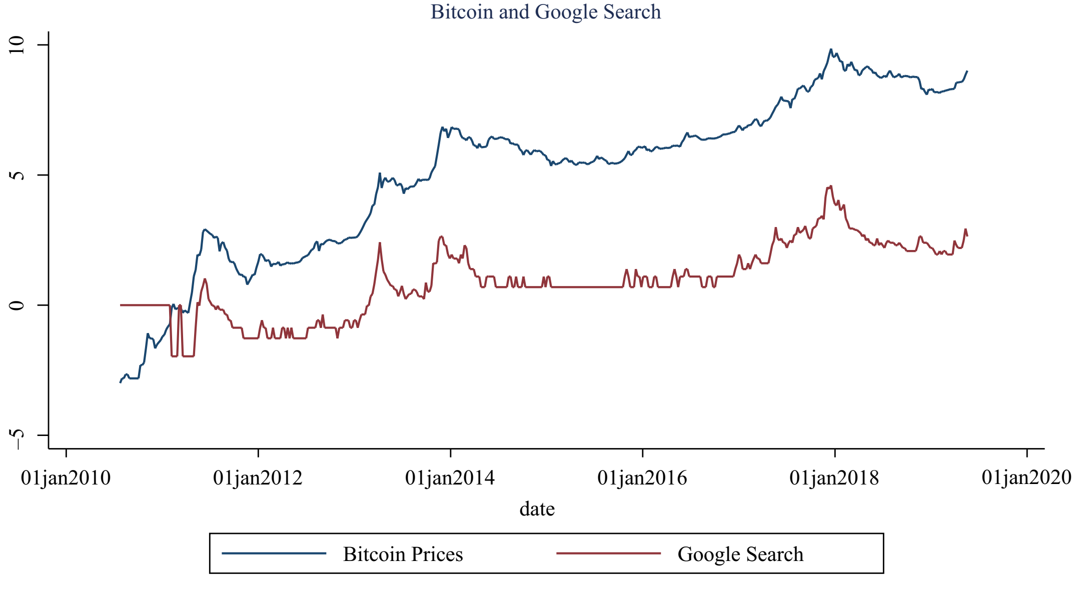

FIGURE A1 Dynamics of the Bitcoin prices and Google search associated with Bitcoin from July 2010 to May 2019. Prices are in log terms [Colour figure can be viewed at wileyonlinelibrary.com]

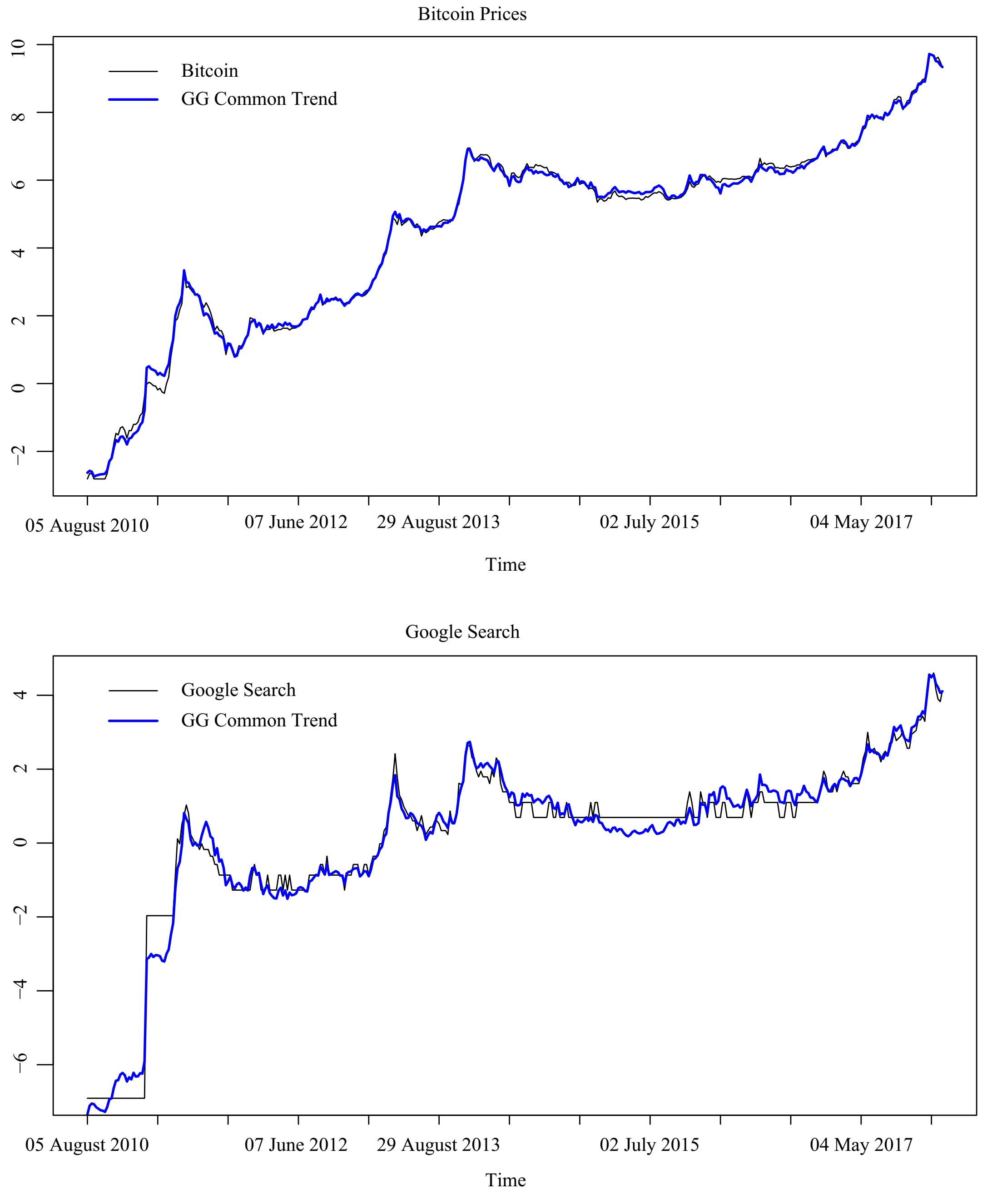

FIGURE A2 Dynamics of the pair (Y_t^*, Y_t) for the time series of Bitcoin and Google search using decomposition (7). The time series Y_t^* are obtained from the system with four common factors. The sample spans the period from July 2010 to January 2018 [Colour figure can be viewed at wileyonlinelibrary.com]

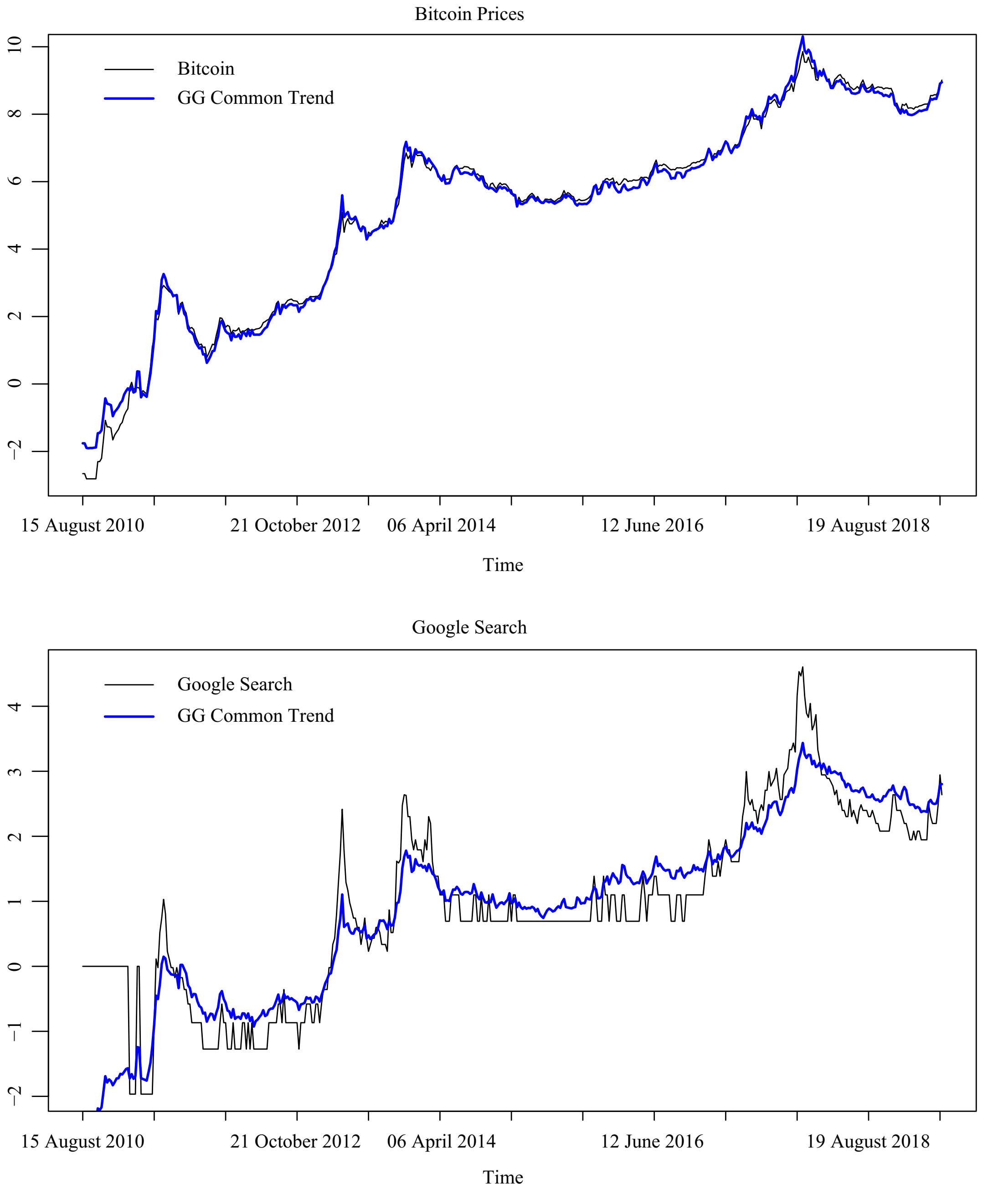

FIGURE A3 Dynamics of the pair (Y_t^*, Y_t) for the time series of Bitcoin and Google Search using decomposition (7). The time series Y_t^* are obtained from the system with four common factors. The sample spans the period from July 2010 to May 2019 [Colour figure can be viewed at wileyonlinelibrary.com]