

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Süssmuth, Bernd

Article — Published Version The mutual predictability of Bitcoin and web search dynamics

Journal of Forecasting

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Süssmuth, Bernd (2021) : The mutual predictability of Bitcoin and web search dynamics, Journal of Forecasting, ISSN 1099-131X, Wiley Periodicals, Inc., Hoboken, USA, Vol. 41, Iss. 3, pp. 435-454, https://doi.org/10.1002/for.2819

This Version is available at: https://hdl.handle.net/10419/264510

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



http://creativecommons.org/licenses/by-nc/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

Revised: 16 June 2021

RESEARCH ARTICLE

WILEY

The mutual predictability of Bitcoin and web search dynamics

Bernd Süssmuth^{1,2}

¹University of Leipzig, Leipzig, Germany ²CESifo

Correspondence

Bernd Süssmuth, Institute for Empirical Research in Economics, Grimmaische Str. 12, 04109 Leipzig, Germany. Email: suessmuth@wifa.uni-leipzig.de

Abstract

Economic theory predicts the price dynamics of an unbacked asset to be inherently unforecastable. The same applies to exchange rates of unbacked currencies. Albeit, empirically investors are found to be driven by online and offline news media. This study analyzes the Bitcoin cryptocurrency price series and web search queries with regard to their mutual predictability and causeeffect delay structure. Chinese Baidu engine searches and compounded Baidu-Google search statistics predict Bitcoin price dynamics at relatively high frequencies ranging from 2 to 5 months. In the other direction, Grangercausality runs from the cryptocurrency price to queries statistics across nearly all frequencies. In both directions, the reaction time computed from a phase delay measure for the relevant frequency bands with significant causality ranges from about 1 to 4 months. For either direction, out-of-sample forecasts are more accurate than forecasts of a benchmark stochastic process. Bivariate models including the Baidu Search Index slightly outperform competing models that include a Baidu-Google composite index. Predictive power seems less diluted if the September 2017 trade regulations by the Chinese government are controlled for.

K E Y W O R D S Bitcoin, bubbles, causality, frequency domain

1 | INTRODUCTION

Terms like cryptocurrency, crypto-token, and blockchain technology are on everyone's lips. Most prominently, this applies to Bitcoin (BTC); see Cheah and Fry (2015) and Bariviera et al. (2017). The currently high-pitched online as well as offline media interest can be interpreted as a hype in the sense of an overintensified publicity or promotion of the BTC. In February 2021, the supposedly richest man on earth at the time, Elon Musk, by communicating his support for the crypto-token through his social media channels, let the BTC price skyrocket to an all time high in the first few weeks of the trading year. There is a lot of narrative on the relationship between popularity in the form of news coverage—in particular, conveyed through social media—and BTC trading as well as pricing. Prominent in this context is, for instance, the case of Nasdaq-listed Long Island Iced Tea Corp (LTEA) in the last quarters of 2017. In the third quarter of 2017, LTEA faced a net loss amounting to US dollars (USD) 3.9 million. When LTEA announced its realignment of business and its renaming into Long Blockchain Corp

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. © 2021 The Author. *Journal of Forecasting* published by John Wiley & Sons Ltd. 436 WILEY-

(LBCC) on December 21, 2017, its shares soared and tripled in value within hours; see Figure 1. To the present, LBCC continues a non-alcoholic beverage subsidiary.

This zeitgeisty phenomenon motivates the general research question of the present study: Is the allocation of funds to virtual currencies or risky assets such as the BTC increasing with public interest, attention, and popularity of the financial vehicle or its publicity in general? Or is it the other way around? Taking into account query statistics of the dominant search engine in China, the present study investigates the mutual predictability of BTC price dynamics and attention measures. The latter are chosen to be more adequate and comprehensive than the ones previously used in the literature. At first, this is done in an in-sample analysis. The general finding is that the BTC price is helpful in predicting the hype more or less immediately. However, search data are still a good predictor at frequencies with intra-one-quarter periodicities and a delay of about 2 months giving room for pursuing anticyclical investment strategies. In a second part of the analysis, this is generally also confirmed in an outof-sample predictive power assessment.

The overarching research question in a more narrow sense is to quantitatively assess how helpful internet data from "secondary sources" (Edelman, 2012) are in predicting and, ultimately, in modeling and out-of-sample forecasting Fintech-related price dynamics at the beginning of the 2020s. The existing literature does not provide clear-cut answers. On the one hand, theory (e.g., Manuelli & Peck, 1990.) predicts price dynamics of an unbacked asset or currency to be inherently unforecastable. On the other hand, recent models find it, irrespective of fundamentals, optimal for investors to delegate the collection of information to news media when facing attention constraints (Nimark & Pitschner, 2019). That news holds the potential to drive the performance of an asset is empirically confirmed, among others, by Solomon et al. (2014). In particular, in the context of cryptocurrencies that are traded online, it is highly probable that news, in turn, is driven by web searches. To some extent, this relates the present paper to empirical models of contemporary finance. See, for example, Antweiler and Frank (2004) and Chen et al. (2014) who study online and offline traded stocks and their interaction with internet message boards and social media attention. However, the perspective chosen here is macroeconomic or "macrofinance" with a focus on predictability, cause-effect structure, and future theoretical modeling relying on a monthly observation frequency (e.g., Jermann, 2018.).

The methods of the present study comprise a battery of multivariate time series techniques in the time and frequency domain such as testing for Granger-causality in the frequency domain (Breitung & Candelon, 2006) and forecast accuracy testing between models in the time domain (Diebold & Mariano, 1995; Harvey et al., 1997). They have been successfully applied in diverse contexts; see, among many others, Tastan (2015), Aluko and Adeyeye (2020), and Fromentin and Tadjeddine (2020) for spectral Granger-causality applications and Carstensen et al. (2011) and Duarte and Süssmuth (2018) for applications of forecasting ability tests across models. To assess the cause– effect reaction time for frequency bands with significant Granger-causality, a phase delay measure recently developed by Breitung and Schreiber (2018) is computed.

2 | LITERATURE REVIEW

The related literature to analyze the sketched phenomenon is growing and predominantly, though not



FIGURE 1 Nasdaq US dollars (USD) opening value of LTEA (LBCC) stock, 2017. LBCC, Long Blockchain Corp; LTEA, Long Island Iced Tea Corp

exclusively, published in economics journals. It comprises, among others, the survey by Böhme et al. (2015), the studies by Cheah and Fry (2015), Urguhart (2017), and Akyildirim et al. (2021) treating BTC as financial asset prone to speculative bubbles and price clustering, and Brandvold et al. (2015) as well as Gandal et al. (2018) treating BTC as globally exchange-traded currency. The most recent study by Akvildirim et al. (2021) comes close to the present one by its focus on predictability. In contrast to the approach chosen here, Akyildirim et al. (2021) analyze not only the price dynamics of the BTC but also of the other 11 most liquid contemporary cryptocurrencies. What is more, in a horse race fashion, the relative performance of a battery of four different machine learning algorithms is comprehensively investigated. Other seminal interdisciplinary contributions emanating either from the information systems or the econophysics field of study include Garcia et al. (2014), Bariviera et al. (2015), Kristoufek (2015), Li and Wang (2017), and Alvarez-Ramirez et al. (2018). The pioneering study on web search queries and BTC prices by Kristoufek (2013) referred to BTC as a currency but treated it rather as an asset. Although it is a seminal study, it can be seen as suffering from two crucial drawbacks that the present analysis overcomes. First, it only covers the seed phase or early trading period of the BTC from mid-2011 to mid-2013. Secondly, it relies on web search queries on Google and Wikipedia only. Both were and still are blocked in the People's Republic of China. Thus, they were and still are inaccessible for a substantial share of investors, users, and miners of BTC at the time and to the present.¹ During November and December 2013, for example, roughly half of all BTC trades were made in Chinese yuan (Brandvold et al., 2015, p. 20). Ciaian et al. (2018, p. 178) note regarding the regional distribution and the trading currency composition for the BTC that-while the United States and the USD dominated the BTC market in the first years after its introduction—nowadays "almost all [BTC] trading is done in China." The authors document the "staggering rise of China as the dominant trader" of BTC by showing that from less than a 10% share in January 2012, the yuan made up nearly 100% of all BTC trading by the end of 2016. Although this share declined at the latest since fall 2017, when the Chinese government announced to block the access to foreign ICO and crypto-to-fiat exchanges, it seems fair to state that over the period of analysis of the

present study (from mid-2011 to the first quarter of 2018), the average share is at least 50%. This fact renders Google and Wikipedia series a seriously incomplete measure of attention allocation or (potential) investors' interest.

In contrast, the present study does not only rely on Google Trends Statistics (GTS)-provided in normalized terms within the frequency of choice of one month by Google Inc./LLC-for "Bitcoin" searches as, for example, in Cheah and Fry (2015). Chinese web search engine Baidu non-normalized query statistics for "比特币" (i.e., "Bitcoin") is also used. The Baidu Zhishu or Baidu Search Index (BSI) is the query statistics of the by far most commonly used web search engine in China given the Google ban that preceded the BTC launch by a couple of months. BSI data cannot be directly downloaded, but are accessible on a selective basis and report absolute query figures. The observation frequency of BSI data available for this study is monthly. Its informative content has been recently approved by Liu et al. (2016), Shen et al. (2017), and He et al. (2018) using it to successfully predict dengue fever outbreaks, Chinese stock returns, and HIV incidences in contemporary China, respectively. According to NetMarketShare (netmarketshare.com), tracking usage shares of web technologies, the market shares of Google and Baidu in August 2018 amounted to about 70% and 20% for mobile devices and 76% and 11% for desktop/laptop devices, respectively. None of the other engines even just nears a double digit usage share.

To sum up the crucial point made in the last paragraphs, the previous literature suffers from a substantial deficiency by more or less completely disregarding the regional origin and distribution of BTC-related activity. Ignoring that since mid-2011 about half of BTC trades and investments on average emanate from China lets indicators for attention allocation and investors' curiosity, which are exclusively based on online services such as Google, Wikipedia, or Twitter known as blocked or censored in China, appear as unsatisfactory. This concerns all of the most recent and of the most related studies (Aalborg et al., 2019; Ciaian & Rajcaniova, 2016; Kjærland et al., 2018; Kristoufek, 2013; Garcia et al., 2014).

This study is the first to overcome this deficiency by explicitly considering BSI series in the context of the mutual predictability of BTC dynamics and web search statistics.

3 | DATA

BTC price series of daily frequency are obtained from CoinDesk (https://coindesk.com) denoted in USD. CoinDesk provides the data as Greenwich Mean Time

¹An exemplary more recent study, covering the period from January 2013 to February 2018, in the tradition of Kristoufek (2013) is Kjærland et al. (2018). However, it representatively also suffers from relying exclusively on web queries performed by means of the Google web search engine.

(GMT) end-of-day closing index price, where the latter is aimed to capture the standard retail price reference for industry participants and accounting professionals. It represents an average of leading global BTC exchanges that conform to certain minimum criteria for price discovery and validity.² The historical index data commence on July 1, 2013. Any data prior to that date are based on the Mt. Gox price data (see, e.g., Cheung et al., 2015.). As the frequency of the analysis is monthly-as is, for instance, also the case in Jermann (2018)-monthly averages of daily prices are constructed. Ultimately, the reason why this study uses monthly data is that the provided BSI series are monthly figures. However, it also avoids critical issues of possibly heteroskedastic inter-weekly and intra-weekly regularities in BTC trading. These include, among others, bias due to weekends, secular holidays, and festivities. The conditional heteroskedasticity issue against the backdrop of temporal aggregation will be discussed in somewhat more detail at the end of this section. GTS for search string "Bitcoin"³ can be retrieved in monthly frequency and solely with the maximum value of monthly reported queries automatically normalized to 100 within the sample period. Monthly arithmetic averages of the CoinDesk BTC prices are taken and the resulting series analogously to the GTS series normalized in order to conform prices to normalized query statistics and to avoid to induce spurious cointegration through nonconformity.

To express the CoinDesk BTC monthly price averages in yuan units of Chinese Renminbi (CNY), that is, to measure the BTC/CNY exchange rate in quantity quotation, CNY/USD exchange rate series are obtained at monthly frequency from the online database of the Federal Reserve Bank of St. Louis (FRED) coded as EXCHUS. Based on the argument of market liquidity, that is, "there was practically no liquidity" in the BTC market prior to May 2011 according to the quantitative assessment in Kristoufek (2013, p. 2), the start of the overall sample period is set to May 2011. This concerns both the BTC series and the Baidu Zhishu/Index queries for search string "比特币" (i.e., the Chinese word for "Bitcoin"). The latter is obtained from https://zhishu. baidu.com; see the Appendix A1 for a bit more detail on retrieval.

For reasons of illustration and comparability, one can proceed in analogy to the USD and GTS series and index both series such that the respective monthly maximum value within the sample period corresponds to 100. The overall sample period runs from May 2011 to either January or March 2018. Ending the sample for the mutual BSI queries BTC/CNY exchange rate relationship in January 2018 is justified as by February 2018 the Chinese government announced to block the access to all foreign cryptocurrency exchanges' websites. As the ban took effect during February 2018, it is reasonable to assume that this structural break mostly affected the web search behavior for " $thm \pi$ " in China. Presumably, it induced a confounding downward bias in queries statistics.

Figure 2 shows in its left schedule the GTS series depicted on the left ordinate. The monthly BTC/USD exchange rates in indirect, that is, quantity, quotation, before normalization, is depicted on the right ordinate of this schedule for the period from May 2011 to March 2018. The right schedule displays the normalized BSI series depicted on the left ordinate and the BTC/CNY exchange rate series, also in quantity quotation and before normalization, depicted on the right vertical axis, respectively. The series in this schedule run from May 2011 to January 2018.

There is no reasonable BTC price index or deflator in order to transform nominal BTC exchange rates into real terms. As Foley et al. (2019) find in a recent empirical study identifying darknet marketplaces and combining them with seizures of BTC by law enforcement agencies in their projections, about 25% of all BTC users are engaged in criminal business. Additionally, the study shows that about 50% of all transactions in BTC are related to illegal activities such as drug trade, illegal pornography, and murder-for-hire. As there exist no legal market and corresponding prices for the latter, the construction of an adequate price index or an other reasonable device to express nominal BTC exchange rates in real terms is infeasible. Moreover, to treat BTC as a currency and to ex post date possibly short-lived bubble dynamics as in Cheah and Fry (2015), Cheung et al. (2015), Su et al. (2018), Li et al. (2019), and Hafner (2020) seems unsatisfactory. However, growth rates of historical BTC exchange rate series, expressing the value of the BTC in USD or in CNY, might vary with growth rates of corresponding web search engine queries ("Bitcoin" in case of Google trends and "比特币" in case of the Baidu Zhishu) over time with some negative feedback, such that they can be tested to be cointegrated over longer periods of time.⁴ If then for the queries series an

²See https://www.coindesk.com/coindesk-launches-proprietary-bitcoinprice-index.

³As noted in Kristoufek (2013, p. 3), the reported query frequency is not "case sensitive" in the sense that various search string versions of the word (such as "BitCoin" or "bitcoin") are included.

⁴The idea of feedbacks with digital traces of collective social behavior is, generally, also taken up in quantitative studies emanating from computational linguistics and social and information networks; see, for example, Loughran and MacDonald (2011) and Garcia et al. (2014).



FIGURE 2 Normalized queries and Bitcoin (BTC) exchange rates, 05/2011 to 01/2018 (03/2018). Note: Exchange rates in indirect/ quantity quotation are based on monthly means of daily closing prices. Sources: FRED, CoinDesk, Google, Baidu; Summary statistics: Table 1

	Mean	Std. dev.	Min	Max	N obs	Range
BTC/USD	1182.98	2723.23	2.68	15,065.28	83	05/2011-03/2018
	979.41	2420.78	2.68	15,065.28	81	05/2011-01/2018
EXCHUS	6.39	0.24	6.05	6.92	81	05/2011-01/2018
BTC/CNY	6418.4	15,865.47	17.05	99,328.39	81	05/2011-01/2018
GTS	6.92	13.98	0	100	83	05/2011-03/2018
	6.31	13.53	0	100	81	05/2011-01/2018
BSI	414,064.6	597,153.6	3686	2,794,942	81	05/2011-01/2018
BSI (norm.'ed)	14.81	21.37	0.13	100	81	05/2011-01/2018
B-G-C	0.106	0.153	0.001	0.820	81	05/2011-01/2018

 TABLE 1
 Descriptive statistics of used time series

adequate bubble detection test finds a date preceding the date of the one of the exchange rate series, the impending web searches' bubble burst foreshadows the one in the respective exchange rate series. Hence, if dynamics in BTC exchange rates and web searches share a long-run stochastic trend and the hype measure is related to the price across high and medium term frequencies, shortterm forecasts of the dynamics in web searches may act as an early warning device for a bubble burst in the BTC rates. Given the mere popularity of the latter, which sees many hundreds of millions of USD (or CNY) worth of transactions across its system on a daily basis, let the above sketched strategy look more promising than a strategy solely resting on ex post bubble detection testing. However, at least for the present period of analysis and time series, this strategy turns out as a futile endeavor. This is due to the fact that explosive dynamics in BTC exchange rates can be shown to generally antedate

explosive behavior in web queries for BTC search terms irrespective of any cointegrating relationships. For details, see section 2 in Süssmuth (2019).

In the following, we therefore rather choose to interpret the BTC in a global asset perspective. In fact, there are several recent indications speaking in favor of such a treatment of the BTC, that is, rather as an asset, in the sense of an actual investment or a precursor of an investment,⁵ than as a digital currency. These include the discussion of BTC as part of the underlying technology of near future investment opportunities such as the so-called "internet of payments" for the "internet of things." Corresponding examples comprise future versions of Amazon's "dash button," possible future payment options of Tesla e-vehicles, or any other form of the blockchain economy in general. An indication

⁵That is neither in the narrow sense of a security nor of a commodity.

440 WILEY-

pointing in this direction is the registering of domain names amazonbitcoin.com in March 2013 and amazoncryptocurrency.com in October 2017 by Amazon Technologies Inc.; see, for example, https:// www.whois.com/whois/amazonbitcoin.com. Additionally, there exists a growing number of financial market securities that are based on the BTC as central underlyer such as Australian Apollo Capital Fund launched in February 2018, the Postera fund launched in Liechtenstein in March 2018, or the different products and services provided by Crypto Fund AG approved in 2018 by Switzerland's principal stock exchange (Six Swiss Exchange) and the Swiss Financial Market Supervisory Authority (FINMA). For the manifold future role of BTC and its current role as a pioneering new technology in the financial industry in general, see Yermack (2017).

For treating BTC as a global asset, there is no necessity to separate Google from Baidu queries (i.e., GTS from BSI series). This is due to the fact that BTC price series need not to be analyzed as exchange rate series in either USD or in yuan units of CNY. It also seems straightforward to compound the two query series to a composite index. Even if this strategy would come at the cost of a dilution of the information contained in the raw BSI series, a Baidu–Google composite (B-G-C) index of queries simplifies the analysis.

As argued above, a B-G-C index captures about 90% of total web search engine usage in recent times. Due to a potential structural break in the series induced by Chinese governmental blocking of the access to nondomestic online BTC trading platforms in February 2018, the period of analysis is in the following restricted to run from May 2011 to January 2018. An equal weightening in the construction of the B-G-C index is justified as in the period of observation, on average, about half of BTC trades and investments emanate from China. BSI queries are normalized in the same way as are the retrieved GTS series summed up and divided by 200; see the upper row of diagrams and the lower left diagram in Figure 3. The lower right diagram shows the series in the lower left diagram with a linear trend removed.

A monthly frequency of series may be unconventional—though, not unique; see, for example, Jermann (2018)—in the context of modeling trading data.



FIGURE 3 Constituent and composite queries benchmarked to BTC price, 05/2011–01/2018. Note: Log levels of BTC price, GTS-, BSI-, B-G-C-queries; BTC price, B-G-C-queries cycle components. Sources: CoinDesk, Google, Baidu; Summary statistics: see Table 1

However, it avoids taking into account the issue of conditional heteroskedastic (CH) effects. As shown in Drost and Nijman (1993), low order GARCH processes are not closed in the sense of surviving an increasing sampling interval, that is, the temporal aggregation of underlying time series. If price and popularity series result from aggregating more and more, CH effects disappear. Generally, for example, monthly exchange rate series are found to be homoskedastic while corresponding daily and weekly series are not; see Baillie and Bollerslev (1989). Hafner (2008, p. 476) shows that this result of unclosedness also holds for multivariate-including the following case of bivariate-processes if series represent flow variables. This reasoning lets us refrain from considering CH effects in the following vector autoregressive (VAR) estimations.6

4 | IN-SAMPLE ANALYSIS

Since the dynamics in the focus of this section hinge, as opposed to central parts of the analysis in Süssmuth (2019), not on a long-run equilibrium relationship but rather represent dynamic relationships at business cycle, or even higher, frequencies, cointegration is not an issue.⁷ Our choice of method, that is spectral Granger-causality testing, generally requires stationarity of the series underlying the constituent bivariate models. However, as will be argued in more detail below, for certain specifications in levels of the variables pretesting for integration is obsolete (Breitung & Candelon, 2006; Dolado & Lütkepohl, 1996 and Toda & Yamamoto, 1995).

A standard VAR process of order p, that is, a VAR (p), generating two series x_t and y_t , in reduced form will serve as our methodological starting point. That is

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = Dd_t + \sum_{i=1}^p \begin{pmatrix} a_{11,i} & a_{12,i} \\ a_{21,i} & a_{22,i} \end{pmatrix} \begin{pmatrix} x_{t-i} \\ y_{t-i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{pmatrix}, \quad (1)$$

⁷Readers interested in a comprehensive cointegration analysis, including bubble detection testing based on the methodology proposed by Phillips et al. (2011) and Homm and Breitung (2012), of the series at stake are referred to section 2 in Süssmuth (2019). where d_t denotes a deterministic vector containing constants, trend component, and possibly other exogenous variables,⁸ $a_{kl, i}$ denote coefficients and $[\varepsilon_{xt}, \varepsilon_{yt}]'$ serially uncorrelated reduced form errors.

In this bivariate system, Granger-causality or predictability is defined as follows. There is no Granger-causality given if the prediction of *x* is not improved by lagged values of *y*. As *y* does not help to predict *x*, it is not "Granger-causing" *y*. Thus, if $a_{12,i} = 0$ for i = 1, ..., p, y is found to be not Granger-causal for *x*. Therefore, the adequate choice of tests stems from the F/χ^2 class and, as applied in block-form, represents Wald tests. In general, cointegration implies Granger-causality but not the other way around.

4.1 | Spectral Granger-causality tests

In the following, Breitung–Candelon–Granger- (BCG-) causality testing in the frequency domain (Breitung and Candelon, 2012) is considered as a more adequate method in the context of the present study. This is due to its potential to continuously quantify mutual predictability at classical business cycle and higher frequencies.

Any *n*-dimensional stationary process \mathbf{Y}_t has a spectral representation at frequencies $\omega \in [-\pi,\pi]$ in the form of a spectral density matrix $\mathbf{F}(\omega)$. It is given by the Fourier transform of the covariance function $\gamma_{jk}(\tau)$, $\tau = 0, \pm 1, \pm 2, ...,$ for all j = 1, ..., n; k = 1, ..., n. As $\mathbf{F}(\omega)$ is even, it is sufficient to examine it in $[0,\pi]$.

It can be written as

$$\mathbf{F}(\omega) = \frac{1}{2\pi} \sum_{\tau = -\infty}^{+\infty} \mathbf{G}(\tau) e^{-i\omega\tau}, -\pi \le \omega \le \pi \text{ with} \qquad (2)$$

⁸The recent findings of Kjærland et al. (2018) speak in favor of a parsimonious specification of these exogenous factors as these authors identify no significant impact neither from the technological factor "hashrate" (i.e., the variable speed at which a computer can complete an operation in the BTC code) nor from oil or gold. Additionally and in contrast to autoregressive terms and publicity measured by-in the present study also allowed to be acting as endogenous-Google web search queries, a popular volatility index (VIX) and the BTC transaction volume turn out throughout their specifications as statistically insignificant exogenous determinants of BTC return dynamics. Jermann (2018, p. 6) notes that fluctuations in the BTC price have been essentially uncorrelated with stock markets (empirically also justifying the abstraction from excess returns and risk premiums). For the predictability of a range of BTC risk measures, using high-frequency data, the reader is referred to the comprehensive study by Trucíos (2019).

⁶In this context, it is noteworthy that Jermann (2018) recently tests a model of hyperdeflation for the BTC price also relying on time series that are averaged to a monthly period and abstracting from modeling CH.

$$\mathbf{G}(\omega) = \begin{pmatrix} \gamma_{11}(\omega) & \dots & \gamma_{1n}(\omega) \\ \vdots & \ddots & \vdots \\ \gamma_{n1}(\omega) & \dots & \gamma_{nn}(\omega) \end{pmatrix} \text{ and }$$
$$\mathbf{F}(\omega) = \begin{pmatrix} f_{11}(\omega) & \dots & f_{1n}(\omega) \\ \vdots & \ddots & \vdots \\ f_{n1}(\omega) & \dots & f_{nn}(\omega) \end{pmatrix}.$$

An implementation of Equation (2) can be achieved by a VAR estimation with coefficient matrix **A**. In this case,

$$\mathbf{F}(\omega) = \frac{1}{2\pi} \mathbf{A}(\omega)^{-1} \boldsymbol{\Sigma} \mathbf{A}(\omega)^{-*}, \text{ where}$$
(3)

 Σ is the positive-definite covariance matrix of errors; $A(\omega)$ denotes the Fourier transform of matrix lag polynomial A(L) and "*" its conjugate complex transpose, respectively. According to Wold's theorem

$$\boldsymbol{\Gamma}(L)\mathbf{Y}_{t} = \varepsilon_{t} \Leftrightarrow \begin{pmatrix} x_{t} \\ y_{t} \end{pmatrix} = \Psi(L) \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \end{pmatrix}, \text{ where }$$
(4)

 η_{jt} denotes Choleski factorized errors and $\Psi(L) = \tilde{\Gamma}^{-1}$. Geweke (1982) is the first to make use of this property by equating a respective measure of linear feedback to zero under the null in a test of linear dependence and feedback between multiple time series. That is, for instance, for testing for linear feedback to run from *y* to *x* across all $\omega \in [0, \pi]$

$$H_0: M_{y \to x}(\omega) = \log \left(1 + \frac{|\psi_{12}(e^{-i\omega})|^2}{|\psi_{11}(e^{-i\omega})|^2} \right) = 0, \text{ where } (5)$$

 $y \rightarrow x$ denotes y is helpful in predicting x.

Analogously, one may rewrite (1) using lag polynomial notation as

$$\begin{bmatrix} a_{11}(L) & a_{12}(L) \\ a_{21}(L) & a_{22}(L) \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{bmatrix}.$$
 (6)

In this notation, x_t is not Granger-causal for y_t if $H_0: a_{21}(L) = 0$. To test this hypothesis, the second equation of system (6), that is,

$$\begin{aligned} y_t &= \alpha_1 y_{t-1} + \ldots + \alpha_p y_{t-p} + \beta_1 x_{t-1} + \ldots + \beta_p x_{t-p} + \varepsilon_{yt} \\ &= \alpha(L) y_{t-1} + \beta(L) x_{t-1} + \varepsilon_{yt}, \end{aligned}$$
 (7)

where $\alpha_i = a_{22,i}$ and $\beta_i = -a_{21,i}$, is used. Following Breitung and Candelon (2006), the definition of the falsi-fied BCG-causality under the null then reads as follows.

Definition 1. x_t is not a cause of y_t at frequency ω if the gain function of the filter $\beta(L)$ equals zero at frequency ω , that is,

$$H_0: \left|\beta e^{\omega i}\right| = \left|\sum_{j=1}^p \beta_j \cos(j\omega) + \sum_{j=1}^p \beta_j \sin(j\omega)i\right| = 0.$$

The necessary and sufficient conditions for the inexistence of BCG-causality running from x_t to y_t , thus, are $\sum_{j=1}^{p} \beta_j \cos(j\omega) = 0 \wedge \sum_{j=1}^{p} \beta_j \sin(j\omega) = 0$. The usual *F*-test logic for a linear combination of $R(\omega)\beta = 0$ applies.

For the standard representation of a VAR(p) extended to a VARX(p), as shown in (1) above, it is straightforward to extend the framework of spectral Granger-causality to the case of an additional exogenous variable z, that is, typically to the case of a deterministic trend function

$$x_{t} = c_{1} + \sum_{j=1}^{p} \alpha_{j} x_{t-j} + \sum_{j=1}^{p} \beta_{j} y_{t-j} + \sum_{j=1}^{p} \delta_{j} z_{t-j} + e_{t}.$$
 (8)

It turns null hypothesis (16) into a conditional one, that is, into H_0 : $M_{y \to x|z}(\omega) = 0$ (Tastan, 2015, pp. 1159– 1160). A mere eyeballing of time series shown in Figure 2 is suggestive for the BTC as well as the queries level series to follow an exponential trend. For natural log expressions of the corresponding levels series (Figure 3), it thus appears appropriate to test for a linear trend to be considered in z. The VAR approach generally presupposes stationarity of included series or the use of firstdifferenced I(1) variables requiring a comprehensive range of pretests for cointegrating relationships. The modified Wald test for Granger-causality proposed by Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) does not necessitate pretesting for cointegration. These authors suggest that the usual Wald statistics will be valid for Granger-causality on p lags of a variable in an overfitted $VAR(p+d_{max})$ with d_{max} denoting the highest order of integration in the system of possibly integrated processes. Breitung and Candelon (2006) propose this strategy to be also followed in the present frequency domain context. Assuming that $d_{\text{max}} > 0$, the corresponding test regression is written as

$$x_{t} = c_{1} + \sum_{j=1}^{p} \left(\alpha_{j} x_{t-j} + \beta_{j} y_{t-j} \right) + \sum_{k=p+1}^{p+d_{\max}} (\alpha_{k} x_{t-k} + \beta_{k} y_{t-k}) + e_{t},$$
(9)

where null hypothesis $M_{y\to x}(\omega) = 0$ involves β_j for j = 1, ..., p only. It can be tested using the standard Wald statistics. As the coefficients on the additional lagged variables are not included in the computation of the test statistics, the same χ^2 -distributed Wald statistics can be used as in the case without lag-augmentation. The same applies to the VARX $(p + d_{max})$ model

$$x_{t} = c_{1} + \sum_{j=1}^{p} \left(\alpha_{j} x_{t-j} + \beta_{j} y_{t-j} + \delta_{j} z_{t-j} \right) + \sum_{k=p+1}^{k=p+1} (\alpha_{k} x_{t-k} + \beta_{k} y_{t-k}) + e_{t},$$
(10)

where *x* and *y* series represent natural log expressions of levels, *z* denotes a deterministic linear trend, and H_0 : $M_{y\to x|z}(\omega) = 0$, that we will also use in the following to test for spectral Granger-causality.

4.2 | Spectral delay assessment

The diagonal elements $f_{11}(\omega), ..., f_{nn}(\omega)$ of $\mathbf{F}(\omega)$ in (2) are the real valued autospectra or power spectra. The off-diagonal elements represent cross spectra $f_{jk}(\omega) = c_{jk}(\omega) - iq_{jk}(\omega)$, consisting of $c_{jk}(\omega)$ cospectra and $q_{jk}(\omega)$ quadrature spectra. The cross-spectra are complex valued functions in ω , but simple manipulations yield the more readily interpretable real phase shift measure

$$\phi_{jk}(\omega) = \arctan \frac{-q_{jk}(\omega)}{c_{jk}(\omega)}.$$
(11)

Both in the autospectral and in the pairwise bivariate case, $\phi(\omega)$ and $\phi_{jk}(\omega) = \phi_{xy}(\omega)$, the phase shift can be visualized either on circular scale or on linear scale or on both scales at a time. For the latter, see, for example, Heer and Süssmuth (2013, p. 406). As the phase shift corresponds to the phase angle –also referred to as angular coordinate or polar angle– in the circular space, it repeats every 2π periods due to the circular diameter equaling 2π . It is, thus, said to be only defined mod 2π . Additionally, as illustrated in detail in Heer and Süssmuth (2013), there is a frequency, where the phase shift reaches π , that is, where the counterclockwisely rotated phase angle coincides with the horizontal originating from the circle center. It is this phase shift that cannot be distinguished

from $\phi_{xy}(\omega) = -\pi$. In this case, there is no difference between the statement "series *y* leads series *x* with half a cycle length" and the statement "series *y* lags series *x* with half a cycle length." A discontinuity at the angular frequency corresponding to this phase shift results.

A solution using the four-quadrant version of the inverse tangent, that is, of the arctan, function that is usually referred to as atan2 is recently proposed by Breitung and Schreiber (2018, p. 64). Reconsider VAR representation (7) above and rearrange, given invertibility of $\alpha(L) = 1 - \sum_{j=1}^{p} \alpha_j L^j$, to obtain

$$y_t = \frac{\beta(L)}{\alpha(L)} x_{t-1} + \nu_t = \rho(L) x_t + \nu_t, \quad (12)$$

where $\rho(L) = \beta(L)L/\alpha(L)$ and $\nu_t = \alpha(L)^{-1}\varepsilon_{yt}$. According to Breitung and Schreiber (2018), the phase shift induced by filter $\rho(L)$, for non-zero gains of implied filters $\beta(L)L$ and $\alpha(L)$ and for the atan2 definition space (0, 2π], is given by

$$\phi_{\rho}(\omega) = \operatorname{atan2}\left(\frac{q_{\rho}(\omega)}{c_{\rho}(\omega)}, \operatorname{sgn}\left[q_{\rho}(\omega)\right], \operatorname{sgn}\left[c_{\rho}(\omega)\right]\right), \quad (13)$$

where, for two general arguments (a, b), atan2 in terms of the standard arctan function can be expressed as follows:

$$\operatorname{atan2}(a,b) = \begin{cases} \operatorname{arctan}\left(\frac{a}{b}\right) & \text{if } a > 0, \\ \operatorname{arctan}\left(\frac{a}{b}\right) + 2\pi & \text{if } a < 0 \text{ and } b \ge 0, \\ 2\pi & \text{if } a = 0 \text{ and } b > 0, \\ \operatorname{undefined} & \text{if } a > 0 \text{ and } b = 0. \end{cases}$$
(14)

Using (13) in combination with (14), the spectral delay measure by Breitung and Schreiber (2018) is calculated as

$$d(\omega) = \frac{\phi_{\rho}(\omega)}{\omega},\tag{15}$$

where $\tilde{\phi}_{\rho}(\omega)$ denotes "unwrapped" phase delay. With phase unwrapping, Breitung and Schreiber (2018, pp. 64–65) refer to the following strategy. If at some frequency ω the term $q_{\rho}(\omega)$ switches its sign for $c_{\rho}(\omega) > 0$ the resulting phase shift will show a discontinuous jump down from (or up to) 2π to (or from) an arbitrarily close to zero value. As the phase shift in principle is only identified up to adding integer multiples of 2π , the implied delay function will jump between $2\pi/\omega$ and zero; see (14) in combination with (15). By definition, however, ₩ILEY-

all phase shifts measured by (13) are to be mapped into the interval $(0, 2\pi]$. The "unwrapping" workaround is to remove discontinuities in the phase shift function by adding or subtracting integer multiples of 2π where needed. As the unwrapping procedure is independent of the estimation of $\phi_{\rho}(\omega)$, it does not bias the sampling uncertainty of locally identified measures.

In analogy to the general requirement of the underlying VAR in the preceding subsection to be of order $p \ge 2$, a VAR order in excess of two lags, that is, $p \ge 3$, is necessary for estimating spectral delay; see Breitung and Schreiber (2018).

4.3 | Findings and interpretation

As argued above, the starting point of the analysis is the estimation of a bivariate VAR(p) system as specified by Equations (1), (4) and (6), (7) considering a linear trend as exogenous variable as captured by Dd_t in Equation (1) or analogously by z in Equation (8). Throughout, the natural log transformed BTC price series in USD serves as x_t

series. For y_t , let us consider first natural log expressions of absolute BSI queries (dashed series in the upper left window of Figure 3) and as an alternative In-transformed B-G-C series (dashed series in the lower left schedule of Figure 3), which by construction are logs of a normalized time series, that is, logs of values from the (0, 1) interval. All estimated VAR models satisfy the stability condition. Additionally, the lag order of the respective VARX model that minimizes for both considered y_t series the AIC and final prediction error (FPE) information criterion is p = 3. A standard F-test finds the linear trend as exogenous variable in the VARX(3) to be significantly different from zero with a p value of 0.04 for BSI queries and a p value of 0.05 for the B-G-C composite series, respectively. The outcome of BCG-causality tests, reported and visualized in Figures 4 and 5 as well as corresponding phase delay measures shown in Figures 6 and 7, rely on estimates of corresponding bivariate VARX(3) models.

To correctly interpret the diagrams in the four figures, it is important to note that frequency depicted on the respective abscissa in the respective four schedules refers



FIGURE 4 Spectral Granger-causality tests using BSI queries as hype measure. Note: First row of diagrams: standard BCG tests; second row: Toda–Yamamoto modified BCG tests; left column of diagrams: price-to-hype (BTC \rightarrow BSI), right column: hype-to-price (BSI \rightarrow BTC). BSI, Baidu Search Index; BTC, Bitcoin



FIGURE 5 Spectral Granger-causality tests using B-G-C queries as hype measure. Note: First row of diagrams: standard BCG tests; second row: Toda–Yamamoto modified BCG tests; left column of diagrams: price-to-hype (BTC \rightarrow B-G-C), right column: hype-to-price (B-G-C \rightarrow BTC). B-G-C, Baidu–Google composite; BTC, Bitcoin

to angular frequency. Therefore, the highest measurable frequency is $\omega^{\max} = \pi = 3.1415$...corresponding to an ordinary frequency of $f^{\max} = \frac{\omega^{\max}}{2\pi} = \frac{\pi}{2\pi} = 0.5$. It is referred to as Nyquist frequency and represents the lowest discernible periodicity of a contained cyclic mode $P^{\min} = (f^{\max})^{-1} = 2$, that is, a two-period (2 months) cyclicality. Additionally, note that the dashed confidence bands displayed in Figures 6 and 7 are computed for a 5% level of significance.

For standard BCG testing, the BTC price is found to be helpful in predicting BSI queries across all frequencies (upper left diagram in Figure 4). Relying on the lagaugmented (Toda–Yamamoto modified) BCG test, shown in the lower left diagram of Figure 4, this is also the case apart from the (0.51, 1.38) frequency band corresponding to periodicities of 4.6 to 12.3 months for the 5% level of significance and from the (0.63, 1.21) frequency band corresponding to periodicities of 5.2 to 10 months for the 10% level of significance. See the four thresholds given by dashed vertical lines in the lower left diagram identifying the respective bands in Figure 4. Combining this insight with results on spectral delay reported in the first column of Figure 6 reveals that the corresponding delay between cause (BTC price) and effect (BSI queries) is significantly different from zero for periodicities lower than 6 and 4 months (corresponding to abscissa values of about 1 and 1.5, respectively). It is of relatively low magnitude, that is, it amounts to just about 1 month.

Regarding the second column of diagrams in Figures 4 and 6, it can be noted that the hype measure BSI queries is Granger-causal for the BTC price series only for frequencies (periodicities) in excess of 1.15 (below 5.5 months) for the 10% level of significance⁹ and standard BCG testing with a delay of about 3 to 4 months.

For the modified BCG test result, there is no Grangercausality found between BSI queries and BTC price that is significant at the 5% level of significance (lower right schedule in Figure 4). For the 10% level, the threshold

⁹At the 5% level of significance, the frequency (periodicity) threshold is 1.36 (3.5 months).



FIGURE 6 Spectral delay estimates using BSI queries as hype measure. Note: First row of diagrams: based on standard bivariate VAR; second row: lag-augmented VAR based; left column of diagrams: price-to-hype (BTC \rightarrow BSI), right column: hype-to-price (BSI \rightarrow BTC). BSI, Baidu Search Index; BTC, Bitcoin

resembles the one of the one found by means of standard BCG testing: 1.22 (5.2 months). The same holds for the corresponding phase delay; see the lower right schedule in Figure 6.

446

 \perp WILEY

For the composite B-G-C series as hype measure, results are visualized in Figures 5 and 7. Focusing on the respective first column of diagrams at first, it can be stated that the BTC price granger-causes the B-G-C hype measure for all frequencies/periodicities, even at a conservative 5% level of significance with a delay of \leq 1 month. The delay is significantly different from 0 at a 5% level for cyclic components with period lengths implying a peak-to-peak distance that is not exceeding half a year or, at least, is lower than 6 months.

Turning to the respective second column of diagrams in Figures 5 and 7, the results for the BSI attention measure are generally also confirmed for the B-G-C queries inasmuch as the hype is causing the price significantly only at high frequencies corresponding to fluctuations with periodicities lower than 4 to 5 months. Also in accordance with preceding test results, the phase delay at relevant frequencies amounts to about 2 to 4 months.

To sum up, although the price in general is helpful in predicting ("driving") the hype more or less immediately,

that is, delayed at maximum by about 1 month, the hype drives the price only for relatively high-frequency dynamics and with a two- to four-times higher delay.

5 | OUT-OF-SAMPLE ANALYSIS

For our out-of-sample predictive power assessment, we will focus-as is common practice-on rather highfrequency (one-period/month ahead) forecasts, where according to our in-sample analysis mutual predictability of price and hype measure series should be given. In order to assess the one-period ahead forecast properties, the period starting May 2011 (2011m5) to July 2015 (2015m7) is treated as given and a recursive window is successively extended by an additional monthly observation to make 1-month ahead projections for each successive step of prolonging the in-sample-window with a length of 50 months up to the penultimate month of the sample. Hence, our first projection uses Equation (1), or analogously the corresponding version of Equation (7), to make a respective forecast for h = 1, ..., 29 following months until January 2018 (2018m1). This procedure is repeated for T_2 (2015m8) to T_{30} (2018m1) in order to obtain a series of one-step ahead forecasts $x_{T_1+1}^F, ..., x_{T_{2n+1}}^F$



FIGURE 7 Spectral delay estimates using B-G-C queries as hype measure. Note: First row of diagrams: based on standard bivariate VAR; second row: lag-augmented VAR based; left column of diagrams: price-to-hype (BTC \rightarrow B-G-C), right column: hype-to-price (B-G-C \rightarrow BTC). B-G-C, Baidu–Google composite; BTC, Bitcoin

or likewise $y_{T_1+1}^F$,..., $y_{T_{29}+1}^F$. Ultimately, we obtain a series of one-step ahead forecasts for T_2 (2015m8) to T_{30} (2018m1). Thus, each one-period ahead forecast series spans 29 months, that is, the 5 months from August 2015 to December 2015 plus the 24 months from January 2016 to January 2018. As a benchmark predictor serves an AR(2) with deterministic trend component.

From $x_{T_1+1}^F, ..., x_{T_{29}+1}^F$ and $y_{T_1+1}^F, ..., y_{T_{29}+1}^F$ for y_t and x_t respectively acting as Granger-causal predictor (or second variable) in the corresponding bivariate VAR (see Section 4), a set of forecast errors (FE) can be computed. The latter includes squared forecast errors (SFE) that put a stronger weight on larger deviations of forecasts from actual values. Additionally, relative FE relate the FE of the respective bivariate forecasting model to the corresponding ones of the AR(2) predictor. A mean relative FE <1 implies, on average, the outperformance of the benchmark model. Mean FE across the VAR models, as described and analyzed in-sample in Section 4, are summarized for the sketched out-of-sample assessment in Table 2. The first two subsets of performance measures, (A1) and (A2), refer to models for which the BSI has been used as hype measure series. For the (B1) and (B2) subsets, the B-G-C series, merging the Google and Baidu engine information, has been used.

At first, we can note that all VAR models outperform the AR(2) benchmark projections as indicated by mean relative FE <1 throughout. In line with the in-sample analysis, both the hype measure and the price series generally also show a mutual predictive power in the one-period ahead out-of-sample forecasting exercise. On average, the popularity measures even show a better mean performance in this high-frequency forecasting range than do the log BTC price series as expressed by mean forecasting errors (FE and SFE) being raised by roughly a factor of 10 comparing the (A1) and (B1) with the corresponding (A2) and (B2) row entries of Table 2, respectively. The difference between the two popularity measures, BSI and B-G-C, is quite slim. It indicates that the predictive power contained in the BSI is, in line with our prior of the substantial role of Chinese BTC investors in the period of analysis, dominating. The mean relative FE are throughout smaller in the case of the BSI series acting as the hype measure. The lowest relative FE mean of 0.594 is found for the log BSI search series forecasted in a bivariate Toda-Yamamoto modified VAR specification with log BTC prices as second variable (as outlined in Section 4); see the (A2) subset of column 2 in Table 2.

As documented, for instance, by the US Library of Congress (LoC), especially since September 2017 (https://

448 WILEY-

	Standard VAR	Toda-Yamamoto VAR	AR(2)
$(A1) \text{ BSI} \rightarrow \text{ BTC}$			
FE	0.00049	0.00141	-0.00465
SFE	0.02504	0.02490	0.03374
Relative FE	0.64680	0.91456	—
(A2) BTC \rightarrow BSI			
FE	0.08419	0.08011	0.06069
SFE	0.29094	0.27686	0.32159
Relative FE	0.63794	0.59401	—
(B1) B-G-C \rightarrow BTC			
FE	-0.00064	-0.00150	-0.00465
SFE	0.02570	0.02502	0.03374
Relative FE	0.67611	0.97655	—
(B2) BTC \rightarrow B-G-C			
FE	0.05732	0.05290	0.04585
SFE	0.19975	0.19243	0.22796
Relative FE	0.81961	0.76277	_

TABLE 2Mean forecast errorsacross models (unconditional)

<i>Note</i> : VAR models and AR(2) specified in log levels with deterministic linear trend component; $y \rightarrow x$	х
denotes y acting as second predicting variable in VAR model specifications for variable x.	

	Standard VAR	Toda-Yamamoto VAR	AR(2)
$(C1) \text{ BSI} \rightarrow \text{ BTC}$			
FE	-0.00440	-0.00150	-0.00465
SFE	0.02087	0.02125	0.03374
Relative FE	0.58430	0.86115	—
(C2) BTC \rightarrow BSI			
FE	0.02711	0.02312	0.06069
SFE	0.23206	0.22559	0.32159
Relative FE	0.75543	0.72175	—
(D1) B-G-C \rightarrow BTC			
FE	-0.00395	-0.00156	-0.00465
SFE	0.02122	0.02104	0.03374
Relative SFE	0.61444	0.92541	—
(D2) BTC \rightarrow B-G-C			
FE	0.01754	0.00810	0.04585
SFE	0.17715	0.17129	0.22796
Relative FE	0.82698	0.78610	—

TABLE 3Mean forecast errorsacross models with post-2017m9 control

Note: VAR models and AR(2) specified in log levels with deterministic linear trend component; $y \rightarrow x$

denotes *y* acting as second predicting variable in VAR model specifications for variable *x*; all VAR models include a dummy-identifier for the period starting September 2017.

www.loc.gov/law/help/cryptocurrency/china.php), governmental institutions in China have announced and prepared to take a series of regulatory measures to crack down on activities related to cryptocurrencies; see also Section 3 above. According to the LoC, this was mainly due to the concern over financial risks associated with cryptocurrencies, including the BTC.

To take this circumstance into account, the out-of sample assessment documented in Table 2 is redone including in all VAR model specifications a dummy that **FIGURE 8** One-step ahead forecasts with recursive window: hypeto-price (BSI \rightarrow BTC). Note: First panel: based on bivariate lag-augmented (Toda-Yamamoto) VAR with linear trend component (Section 4); for mean forecasting statistics, see (C1), Table 4; legend: blue line – original series, dashed red line – AR(2), solid red line – VAR; second panel: forecast horizon performance (with trend removed). B-G-C, Baidu–Google composite; BSI, Baidu Search Index



identifies the period starting with September 2017 to the end of the forecasting horizon, that is, to January 2018. With everything else equal to the previous one-period ahead forecasting exercise, the figures shown in Table 3 result. Although covering only a putatively short period of the forecasting horizon, this consideration, on average, improves the overall out-of-sample forecasting potential. In particular, two models stand out. First, the VAR specifications in subset (C1) of Table 3, using the BSI as hype measure, show comparatively strong mean performance statistics across all considered scenarios, where the price dynamics is predicted by means of popularity measures. Second, with an average FE of 0.008 and a corresponding avaerage SFE of 0.171 given in column 2 of the (D2) subset of Table 3, the Toda-Yamamoto modified VAR seems to be first best among the considered models, where log BTC prices help in predicting hype measure dynamics. Although, in relative FE terms, the Toda-Yamamoto modified VAR with log BTC prices as second variable and log levels of BSI searches as first more pronouncedly outperforms the AR(2). Corresponding performance statistics are given in the (A2) subset of column 2 in Table 2. Out-of-sample forecasts based on lag-augmented (Toda-Yamamoto) VAR models for (C1) and for (A2) together with (D2) are illustrated in Figures 8 and 9, respectively.

In Table 4, the first (2×2) -block with numerical column-entries shows the test statistics of the pairwise test proposed by Diebold and Mariano (1995) to analyze whether empirical loss differences between competing models are statistically significant. It is generally seen to be the most influential and most widely used test in the

context of macroeconomic forecasting (see Carstensen et al., 2011., p. 87). The second (2×2) -block of numerical column-entries refers to the test statistics of the modified Diebold-Mariano (MDM) test suggested by Harvey et al. (1997) that corrects for small sample bias. It evaluates whether the average loss difference between a pair of models is significantly different from zero. Under the null, the forecast accuracy of two models is identical; see, for example, Carstensen et al. (2011, p. 87) for detail. Note, each entry of Table 4 displays the respective test statistics for the forecasting performance of each model¹⁰ listed in the respective row vis-à-vis its respective benchmark model in the header to each column. For instance, in the upper left corner of Table 4, the combination of row "Standard VAR" and column "AR(2)" amounting to 2.623 (2.580) for the (modified) DM test corresponds to the respective test statistics of a test of whether the mean difference of MSE of model "AR(2)" with MSE of model "Standard VAR" is zero. A positive value implies that the model behind the subtractive part (here "Standard VAR") performs better in terms of forecasting accuracy, that is, under the null of the two-sided test is equal accuracy. Corresponding *p* values are given in parantheses.

As expected given our in-sample assessment suggesting, in particular, the hype measures to show predictive power at high frequencies for the BTC price dynamics—both tests ascribe a significantly higher

¹⁰Note, all bivariate VAR models underlying the DM and MDM tests reported in Table 4 are unconditional, that is, not including a post-2017m9 control. Including the latter generates very similar outcomes. They are available on request from the author.



FIGURE 9 Recursive window forecasts: price-to-hype (BTC \rightarrow BSI and BTC \rightarrow B-G-C). Note: All shown forecasts are based on one-step ahead recursive window extensions; all VAR models specified as bivariate lag-augmented (Toda-Yamamoto) with linear trend (Section 4); upper two panels: hype measure is BSI, VAR excluding post-2017m9 control; lower two panels: hype measure is B-G-C, VAR including post-2017m9 control; for mean forecasting statistics see (A2) in Table 3 for the upper two panels and (D2) in Table 4 for the lower two panels; legend: blue line – original series, dashed red line – AR(2), solid red line – VAR; for every second panel: forecast horizon performance (with trend removed). B-G-C, Baidu–Google composite; BSI, Baidu Search Index; BTC, Bitcoin

forecast accuracy to the VAR models in contrast to the competitor AR(2) irrespective of the hype measure used to predict the price series. See the (A1) and (B1) parts of Table 4.

The other way around, that is, when the BTC price series is used to predict the web search dynamics, only in the lag augmented (Toda–Yamamoto) VAR case with BSI as to-be-forecasted hype measure, the two-variable model is significantly better in forecast accuracy than the AR(2) at a 10% level. This holds, however, only for the standard DM test. See the second row entries of the (A2) part in Table 4. In most of the cases, there is no statistically significant difference between the forecast accuracy of the standard and the lag augmented (Toda–Yamamoto) VAR. The only exception is the (B2) case, where the predictive accuracy of the BTC price

TABLE 4 Forecast accuracy tests between models

	Diebold-Mariano (DM) Test		Modified DM Test	
(A1)	AR(2)	Standard VAR	AR(2)	Standard VAR
Standard VAR	2.623*** (0.008)		2.580** (0.015)	
Toda-Yamamoto VAR	1.951* (0.051)	0.119 (0.905)	1.919* (0.065)	0.117 (0.907)
(A2)	AR(2)	Standard VAR	AR(2)	Standard VAR
Standard VAR	1.354 (0.175)		1.332 (0.193)	
Toda-Yamamoto VAR	1.662* (0.096)	1.265 (0.206)	1.634 (0.113)	1.244 (0.223)
(B1)	AR(2)	Standard VAR	AR(2)	Standard VAR
Standard VAR	2.580*** (0.009)		2.538** (0.016)	
Toda-Yamamoto VAR	1.895* (0.058)	0.478 (0.632)	1.864* (0.072)	0.470 (0.641)
(B2)	AR(2)	Standard VAR †	AR(2)	Standard VAR †
Standard VAR	1.333 (0.182)		1.311 (0.200)	
Toda–Yamamoto VAR^{\dagger}	1.414 (0.157)	1.962** (0.049)	1.391 (0.858)	1.929* (0.063)

Note: (A1) refers to BSI \rightarrow BTC, (A2) to BTC \rightarrow BSI, (B1) to B-G-C \rightarrow BTC, and (B2) to BTC \rightarrow B-G-C, respectively; for each test, MSE is used as loss criterion; a uniform kernel is employed in the calculation of the long-run variance with the exception of the two[†] indexed cases, where a Bartlett kernel is chosen to guarantee estimates to be positive definite; the maximum lag order in the long-run variance calculation is determined by the Schwert criterion

 $p_{\text{max}} = \left[12 \cdot (T/100)^{\frac{1}{2}} \right]$, where T is sample size and rectangular brackets give the nearest integer part of the argument.

p < 0.10. p < 0.05 p < 0.01

for the Baidu–Google composite series is assessed. In this case, the lag augmented model is significantly more accurate in its predictions. This holds for both the DM and MDM test.

6 | CONCLUSION

This study is the first to overcome a substantial deficiency of the previous literature on BTC price dynamics and web search statistics, which by now has more or less completely ignored the regional origin and distribution of BTC-related activity. Since mid-2011, about half of BTC trades and investments on average emanate from China, motivating the use of Chinese Baidu web searches for "比特币" (i.e., "Bitcoin") and an unweighted Baidu-Google queries composite series as a measure of attention or, colloquially, the hype. Economic theory (e.g., Manuelli & Peck, 1990) predicts that the price dynamics of an unbacked asset is inherently unforecastable. The present study does not confirm this view. Chinese Baidu web searches and compounded Baidu-Google queries predict BTC price dynamics at relatively high frequencies corresponding to fluctuations with periodicities lower than 4 to 5 months. The reaction time at relevant frequencies amounts to about 2 to 4 months. A rationalization of this relationship can be seen in momentum strategies (Jegadeesh & Titman, 2001) or endogenously timed herding (Süssmuth, 2002) rooted in short-term motives rather than fundamentals. It gives room for pursuing anticyclical investment strategies, for example,

based on BTC (USD) monthly future contracts through the BakktTM trading platform recently launched and backed by NYSE parent Intercontinental Exchange (https://www.bakkt.com/index). The reaction time of about one quarter might be interpreted as the time needed for web searches to influence online and offline news, which, in turn, can drive BTC price dynamics. Irrespective of this result, the cryptocurrency price is found predictive for queries statistics across nearly all frequencies with a delay of just about 1 month. The outof-sample analysis of this paper generally confirms this mutual predictability feature as for either direction, highfrequency out-of-sample forecasts are more accurate than corresponding ones of a benchmark AR(2). Bivariate models including the BSI slightly outperform competing models that include a Baidu-Google composite index stressing the role of Chinese investors and the corresponding attention measure in the period from mid-2011 to early 2018. Pointing in the same direction, predictive power is slightly enhanced if the September 2017 trade regulations by the Chinese government are controlled for. It remains for future work and longer time series to show that all in-sample findings can be reconciled also with medium term multi-period ahead out-of-sample foreasting exercises.

Getting back to the overarching question of the economic forecasting potential and relevance of internet data raised in Section 1, one may summarize the central insights as follows. Internet data from secondary sources can act quite well in the short term as predictors for BTC price dynamics. This feature suggests to consider the

WILEY

BTC, for example, as a hedge candidate in periods of up to a maximum of half-year intervals. At the same time, attention measures need to be carefully chosen so that the regional origin and distribution of BTC-related activity is taken care of.

ACKNOWLEDGMENTS

⊥Wiley-

Foremost, I gratefully appreciate very helpful comments and suggestions by two anonymous referees that markedly improved the paper. I would also like to thank Jörg Breitung and Helmut Lütkepohl for a host of very valuable comments on an early draft of the paper as well as Pablo Duarte, Frank Simmen, Benjamin Loeper, and Jingjing Lyu for their excellent research assistance. My thanks are also due to Felix Noth, Julian Hackinger, and participants of the TUM Research Seminar in Economics and of the Time Series Session of the VfS Annual Meeting of the German Economic Association for many helpful discussions. The usual disclaimer applies. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

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

ORCID

Bernd Süssmuth D https://orcid.org/0000-0002-6323-6729

REFERENCES

- Aalborg, H. A., Molnár, P., & de Vries, J. E. (2019). What can explain the price, volatility and trading volume of bitcoin? *Finance Research Letters*, 29, 255–265. https://doi.org/10.1016/j. frl.2018.08.010
- Akyildirim, E., Goncu, A., & Sensoy, A. (2021). Prediction of cryptocurrency returns using machine learning. *Annals of Operations Research*, 297, 3–36. https://doi.org/10.1007/s10479-020-03575-y
- Aluko, O. A., & Adeyeye, P. O. (2020). Imports and economic growth in Africa: Testing for granger causality in the frequency domain. *Journal of International Trade & Economic Development*, 29, 850–864. https://doi.org/10.1080/09638199.2020.1751870
- Alvarez-Ramirez, J., Rodriguez, E., & Ibarra-Valdez, C. (2018). Longrange correlations and asymmetry in the bitcoin market. *Physica A*, 492, 948–955. https://doi.org/10.1016/j.physa.2017.11.025
- Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal* of Finance, 59, 1259–1294.
- Baillie, R. T., & Bollerslev, T. (1989). The message in daily exchange rates: A conditional-variance tale. *Journal of Business and Economic Statistics*, 7, 297–305.
- Bariviera, A. F., Basgall, M. J., Hasperué, W., & Naiouf, M. (2017). Some styilzed facts of the bitcoin market. *Physica A*, 484, 82–90. https://doi.org/10.1016/j.physa.2017.04.159

- Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29, 18–35.
- Brandvold, M., Molnár, P., Vagstad, K., & Vagstad, O. C. A. (2015). Price discovery on bitcoin exchanges. *Journal of International Financial Markets, Institutions & Money*, 36, 18–35.
- Breitung, J., & Candelon, B. (2006). Testing for short- and long-run causality: A frequency domain approach. *Journal of Econometrics*, 132, 363–378. https://doi.org/10.1016/j.jeconom.2005.02.004
- Breitung, J., & Schreiber, S. (2018). Assessing causality and delay within a frequency band. *Econometrics and Statistics*, *6*, 57–73. https://doi.org/10.1016/j.ecosta.2017.04.005
- Carstensen, K., Wohlrabe, K., & Ziegler, C. (2011). Predictive ability of business cycle indicators under test. *Journal of Economics* and Statistics, 231, 82–106.
- Cheah, E.-T., & Fry, J. (2015). Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130, 32–36.
- Chen, H., De, P., Hu, Y. J., & Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, 27, 1367–1403. https://doi. org/10.1093/rfs/hhu001
- Cheung, A. W.-K., Roca, E., & Su, J.-J. (2015). Crypto-currency bubbles: An application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox Bitcoin prices. *Applied Economics*, 47, 2348–2358.
- Ciaian, P., Rajcaniova, M., & Kancs, D. A. (2016). The economics of BitCoin price formation. *Applied Economics*, 48, 1799–1815. https://doi.org/10.1080/00036846.2015.1109038
- Ciaian, P., Rajcaniova, M., & Kancs, D. A. (2018). Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions & Money*, 52, 173–195. https://doi.org/10.1016/j.intfin.2017.11.001
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics, 13, 253–265.
- Dolado, J. J., & Lütkepohl, H. (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews*, 15, 369–386. https://doi.org/10.1080/07474939608800362
- Drost, F. C., & Nijman, T. E. (1993). Temporal aggregation of GARCH processes. *Econometrica*, 61, 909–927. https://doi.org/ 10.2307/2951767
- Duarte, P., & Süssmuth, B. (2018). Implementing an approximate dynamic factor model to nowcast GDP using sensitivity analysis. *Journal of Business Cycle Research*, 14, 127–141. https://doi. org/10.1007/s41549-018-0026-0
- Edelman, B. (2012). Using internet data for economic research. Journal of Economic Perspectives, 26, 189–206. https://doi.org/ 10.1257/jep.26.2.189
- Foley, S., Karlsen, J. R., & Putņinš, T. J. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *Review of Financial Studies*, 32, 1798–1853. https://doi.org/10.1093/rfs/hhz015
- Fromentin, V., & Tadjeddine, Y. (2020). Cross-border workers and financial instability: A frequency domain causality analysis applied to the Luxembourg financial Centre. *Applied Economics Letters*, 27, 280–285. https://doi.org/10.1080/13504851.2019. 1613496
- Gandal, N., Hamrick, J. T., Moore, T., & Oberman, T. (2018). Price manipulation in the bitcoin ecosystem. *Journal of Monetary Eco*nomics, 95, 86–96. https://doi.org/10.1016/j.jmoneco.2017.12.004

- Garcia, D., Tessone, C. J., Mavrodiev, P., & Perony, N. (2014). The digital traces of bubbles: Feedback cycles between socioeconomic signals in the bitcoin economy. *Journal of the Royal Society Interface*, 11, 1–9.
- Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, *74*, 427–431.
- Hafner, C. M. (2008). Temporal aggregation of multivariate GARCH processes. *Journal of Econometrics*, 142, 467–483. https://doi. org/10.1016/j.jeconom.2007.08.001
- Hafner, C. M. (2020). Testing for bubbles in cryptocurrencies with timevarying volatility. *Journal of Financial Econometrics*, 18, 233–249.
- Harvey, D. I., Leybourne, S. J., & Newbold, P. (1997). Testing the equality of prediction mean square errors. *International Journal* of Forecasting, 13, 281–291. https://doi.org/10.1016/S0169-2070 (96)00719-4
- He, G., Chen, Y., Chen, B., Wang, H., Shen, L., Liu, L., Suolang, D., Zhang, B., Ju, G., Zhang, L., Du, S., Jiang, X., Pan, Y., & Min, Z. (2018). Using the Baidu search index to predict the incidence of HIV/AIDS. *Nature (Scientific Reports)*, *8*, 1–10.
- Heer, B., & Süssmuth, B. (2013). Tax bracket creep and its effects on income distribution. *Journal of Macroeconomics*, 38, 393–408. https://doi.org/10.1016/j.jmacro.2013.08.015
- Homm, U., & Breitung, J. (2012). Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics*, 10, 198–231. https://doi.org/10.1093/ jjfinec/nbr009
- Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal* of Finance, 56, 699–720. https://doi.org/10.1111/0022-1082.00342
- Jermann, U. J. (2018). Bitcoin and Cagan's model of hyperinflation, mimeo, Wharton School, University of Pennsylvania.
- 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, 63–81. https://doi. org/10.3390/jrfm11040063
- Kristoufek, L. (2013). BitCoin meets Google trends and Wikipedia: Quantifying the relationship between phenomena of the internet era. Nature (Scientific Reports), 3, 1–7.
- Kristoufek, L. (2015). What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. *PLOS ONE*, 10, 1–15.
- Li, Z.-Z., Tao, R., Su, C. W., & Lobont, O.-R. (2019). Does bitcoin bubble burst? *Quality & Quantity*, 53, 91–105. https://doi.org/ 10.1007/s11135-018-0728-3
- Liu, K., Wang, T., Yang, Z., Huang, X., Milinovich, G. J., Lu, Y., Jing, Q., Xia, Y., Zhao, Z., Yang, Y., Tong, S., Hu, W., & Lu, J. (2016). Using Baidu search index to predict dengue outbreak in China. *Nature (Scientific Reports)*, 6, 1–9.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66, 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x
- Manuelli, R. E., & Peck, J. (1990). Exchange rate volatility in an equilibrium asset pricing model. *International Economic Review*, 31, 83–108.
- Nimark, K. P., & Pitschner, S. (2019). News media and delegated information choice. *Journal of Economic Theory*, 181, 160–196. https://doi.org/10.1016/j.jet.2019.02.001
- Phillips, P. C. B., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?

International Economic Review, 52, 201–226. https://doi.org/10. 1111/j.1468-2354.2010.00625.x

- Shen, D., Zhang, Y., Xiong, X., & Zhang, W. (2017). Baidu index and predictability of Chinese stock returns. *Financial Innovation*, 3(4), 1–8.
- Solomon, D. H., Soltes, E., & Sosyura, D. (2014). Winners in the spotlight: Media coverage of fund holdings as driver of flows. *Journal of Financial Economics*, 113, 53–72. https://doi.org/10. 1016/j.jfineco.2014.02.009
- Su, C.-W., Li, Z.-Z., Tao, R., & Si, D.-K. (2018). Testing for multiple bubbles in bitcoin markets: A generalized sup ADF test. *Japan* and the World Economy, 46, 56–63. https://doi.org/10.1016/j. japwor.2018.03.004
- Süssmuth, B. (2002). Can herding explain cyclic comovement? International Game Theory Review, 4, 33–51. https://doi.org/10. 1142/S0219198902000537
- Süssmuth, B. (2019). Bitcoin and web search query dynamics: Is the Price driving the hype or is the hype driving the Price?, CESifo Working Paper No. 7675
- Tastan, H. (2015). Testing for spectral granger causality. Stata Journal, 15, 1157–1166. https://doi.org/10.1177/ 1536867X1501500411
- Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66, 225–250. https://doi.org/10.1016/0304-4076(94)01616-8
- Trucíos, C. (2019). Forecasting bitcoin risk measures: A robust approach. *International Journal of Forecasting*, *35*, 836–847. https://doi.org/10.1016/j.ijforecast.2019.01.003
- Urquhart, A. (2017). Price clustering in bitcoin. *Economics Letters*, 159, 145–148. https://doi.org/10.1016/j.econlet.2017.07.035
- Yermack, D. (2017). Corporate governance and blockchains. *Review* of *Finance*, 21, 7–31.

AUTHOR BIOGRAPHY

Bernd Süssmuth is a professor of Economics and Econometrics at Leipzig University. He earned his PhD from the University of Munich. He is a research affiliate of the CESifo network. In 2007, he was employed as an assistant professor at UCSB. Further previous affiliations are University of Modena, University of Bamberg, Max Planck Institute for Intellectual Property Munich, TUM Munich University of Technology, and the University of Erlangen Nuremberg. His research focus is on applied econometrics, time series analysis in the time and frequency domain, and forecasting.

How to cite this article: Süssmuth, B. (2022). The mutual predictability of Bitcoin and web search dynamics. *Journal of Forecasting*, *41*(3), 435–454. https://doi.org/10.1002/for.2819

454 WILEY-

APPENDIX

A.1 | Code

To carry out tests and estimations, I partly relied on code and software written by Hüseyin Tastan, Sven Schreiber, and Chrisopher F. Baum. Command suites bcgcausality and dmariano for Stata and function package delayspectral.gfn for open-source econometrics program Gretl have been used.

A.2 | Data retrieval

Chinese BSI data were retrieved on February 19, 2018 and made available to me in monthly frequency. BSI data are directly accessible only by corporate users in China or "practitioners" with "private access"; see Shen et al. (2017, p. 7). The usual disclaimer applies.