A critical investigation of cryptocurrency data and analysis

Article (Accepted Version)

A Critical Investigation of Cryptocurrency Data and Analysis

Carol Alexander* and Michael Dakos
University of Sussex Business School, Falmer, Sussex, BN1 6SL, United Kingdom.

July 24, 2019

Forthcoming in Quantitative Finance. DOI:10.1080/14697688.2019.1641347

Abstract

Market betas of bitcoin relative to a broad crypto market index vary considerably, depending on the data source and the index selected. Even greater differences are found for ether and other cryptocurrencies. An in-depth exploration of the cause of these discrepancies reveals a long-standing incorrect time-stamping of some ranking-site data, and hence also the CRIX market index. Furthermore, individual coin data from some exchanges requires adjusting for unstable prices in the 'stablecoin' tether. Even then, Bitfinex coin prices have de-coupled from prices on other exchanges. Is yet another Bitfinex-tether issue arising? Finally, regarding the risk analysis of coin returns, we argue that this requires highly sophisticated models. But calibrating even the simplest GARCH model is extremely difficult because they are surprisingly sensitive to the data source.

Keywords: Cryptocurrency; Cryptoasset; Price data; Market beta; Markov switching GARCH; Volatility.

JEL classification: C22, C5, E42, F31, G1, G2

*Corresponding author. Email: c.alexander@sussex.ac.uk

Highlights

- Market betas of major coins differ remarkably depending on the data source
- Less than half the papers published since Jan 2017 employ correct data
- Data from some sources have mistakes that require adjusting
- Bitfinex prices diverge from other exchanges even after tether adjustment
- Bitcoin needs complex GARCH models and these are highly sensitive to data source

1 Introduction

It is free and easy to download long historical series on prices of major cryptocurrencies at the daily frequency, indeed even hourly and higher frequency data are available on bitcoin (BTC), ether (ETH), ripple (XRP) and other major coins.[1] It is, therefore, unsurprising that a very large number of empirical studies have appeared during the last few years. Indeed, a quick initial search already shows 124 relevant papers published in academic journals between January 2017 and March 2019, and 28 SSRN relevant discussion papers published in 2018.[2] Unfortunately, over 80 of these use:

1. Data from questionable sources; and/or
2. Non-concurrent time-series data in multivariate analysis; and/or
3. Non-traded prices in portfolio optimisation, efficiency studies, trading strategy development or hedging analysis.

The primary sources for traded price and volume data are the centralised crypto exchanges such as Coinbase, Kraken and Binance. These have an Application Programming Interface (API) service that allows retrieval of a limited history of the order book, and traded prices and volumes, via a variety of data transfer protocols.[3] Historical time series for traded prices and volumes may also be obtained from CoinAPI and Cryptodatadownload, for some of the major coin pairs traded on the most established exchanges.[4] However, most of the academic literature on cryptocurrencies uses non-traded price data on individual coins (or tokens) that are freely downloaded from websites such as Cryptocompare (CC), Coinmarketcap (CM) and Coingecko (CG). These 'coin ranking' companies are so called because they rank both coins and exchanges by trading volume and market capitalisation.[5]

[1]A cryptocurrency (or cryptoasset or 'crypto') is a type of digital asset residing on a blockchain, including coins such as bitcoin, ether and ripple and tokens, which differ from coins because they ride on a non-native blockchain.

[2]That is, a Scopus search on title, abstract or keywords with terms 'bitcoin' or 'cryptocurrency' published in Accounting, Business, Econometrics, Economics, Finance or Management journals yielded 247 papers published between January 2017 – March 2019 and 87 SSRN discussion papers (uploaded between January 2018 – March 2019). From these we judged that 124 published and 28 SSRN papers were on topics relevant to this audience. Many more papers were published in conference proceedings but we do not include those here.

[3]These APIs provide free and open access to each centralised exchange's data, allow the management of trade orders, and interact with users via REST, WebSocket, or FIX protocols. REST is best used for infrequent individual requests; WebSocket provides a constant stream of data; FIX is commonly used for order placement and management. Most exchanges provide libraries in common programming languages that allow access to their API – see e.g. the libraries provided in the Coinbase Pro exchange REST API. The typical output of a price data request to an exchange API is in JSON format – see e.g. the output of a browser-based REST API call to Coinbase Pro that retrieves a list of its products. Note that all of the above holds for centralised exchanges that function outside the blockchain; the executed orders and positions held by investors are recorded internally by the exchange, unless a user explicitly requests the transfer of crypto funds to their own personal wallet or fiat funds to their personal bank account. Centralised exchanges therefore must hold fiat currency and cryptocurrency reserves in order to fulfill any such requests. On the other hand, there also exist (less liquid) decentralised exchanges that operate on smart contract-enabled blockchain platforms such as Ethereum and EOS, and all trade orders are recorded and executed on-chain as smart contract transactions – see Daian et al. (2019) for further details.

[4]CoinAPI is a paid service with limited free access even to individual executed trades and Cryptodatadownload provides some daily, hourly and higher frequency data free.

[5]See www.coingecko.com, www.coinmarketcap.com, and www.cryptocompare.com. Other coin ranking sites include Nomics, Coincall, Coincap, and OnchainFX. The period available depends on liquidity, e.g. bitcoin daily data

A straightforward way to illustrate how empirical results can be influenced by the choice of crypto data is to estimate a simple index regression of the form:

$$r_{it} = \alpha_i + \beta_i R_t + \varepsilon_{it},$$

where r_{it} is the ordinary return on the i^{th} source of the coin price and R_t is the return on the market factor. Table 1 displays the results for daily returns on BTC and ETH, with traded and non-traded prices as dependent variables. Here we examine BTC and ETH traded prices from the Bitfinex, Coinbase, Gemini, Kraken and Poloniex centralised crypto exchanges. Non-traded price indices are obtained from CG, CM and CC.

Table 1: **Market Betas of BTC and ETH w.r.t. CCi30, CRIX and MVDA25 Indices**
Market betas with corresponding t-statistics in parentheses of daily returns on BTC (upper panel) and ETH (lower panel) prices from CG, CM and CC, and from Bitfinex, Coinbase, Gemini, Kraken and Poloniex. The market factor is the return on either the CCi30, the CRIX, or the MVDA25 crypto market index. The sample period is 1 April 2016 – 31 March 2019 for BTC and 1 July 2016 – 31 March 2019 for ETH. Parameters of interest are highlighted in blue and red.

	CG	**CM**	**CC**	**Bitfinex**	**Coinbase**	**Gemini**	**Kraken**	**Poloniex**
				BTC				
CCi30	0.374 (14.6)	0.730 (44.3)	0.744 (44.9)	0.742 (43.4)	0.734 (43.2)	0.743 (44.1)	0.734 (43.5)	0.743 (43.3)
CRIX	0.903 (69.4)	0.528 (20.9)	0.519 (20.1)	0.515 (19.6)	0.506 (19.4)	0.515 (19.7)	0.497 (18.9)	0.521 (19.8)
MVDA25	0.359 (15.4)	0.495 (24.1)	0.509 (24.6)	0.501 (23.7)	0.504 (24.2)	0.508 (24.4)	0.504 (24.3)	0.504 (23.8)
				ETH				
CCi30	0.501 (13.1)	1.008 (37.3)	1.020 (37.4)	1.012 (37.7)	1.012 (36.1)	1.023 (36.0)	0.998 (36.3)	1.021 (37.8)
CRIX	1.041 (33.6)	0.513 (12.0)	0.502 (11.6)	0.489 (11.4)	0.490 (11.2)	0.498 (11.2)	0.471 (10.9)	0.510 (11.8)
MVDA25	0.568 (16.8)	0.763 (25.3)	0.778 (25.6)	0.757 (25.0)	0.778 (25.4)	0.777 (24.8)	0.760 (25.0)	0.772 (25.5)

For the market factor we tested three alternative crypto market indices – the CCi30, the CRIX and the MVDA25 index.[6] The indices include a different number of assets – between 25 and 50 of the largest cap coins, and are structured as follows: the CRIX and MVDA25 indices are cap-weighted, and the assets in CCi30 are weighted by the square root of their market cap.

Due to this difference in the weighting scheme, we expected the β estimates for both BTC and ETH to be smaller with respect to (w.r.t.) CCi30 compared with the other two indices. In fact, this is not at all the case; the correlations between index returns and BTC returns from CC indicate

start as early as April 2013 but ether data begins only in August 2015. Bloomberg and Thomson Reuters also provide aggregated data on coin prices. Data are also not limited to US dollar coin prices; there are indices in other fiat currencies like the euro and Japanese yen, calculated using the same methodology.

[6]See www.cci30.com for CCi30, Trimborn and Härdle (2018) and www.thecrix.de for CRIX, and the MVDA25 methodology.

the opposite – 0.81 for CCi30, 0.52 for CRIX and 0.6 for MVDA25. Moreover, given that ETH is significantly more volatile than BTC we expected the same relationship to hold for their estimated betas, but this is not always the case. So we expected the betas estimated from our simple index model to depend on the choice of index to some extent, but our results are extreme. The first inexplicable observation is that beta estimates for the Coingecko (CG) data are radically different compared with the other data sources. For instance, the Coingecko BTC beta w.r.t. the CRIX is 0.903 while the estimated CRIX betas based on all other BTC sources are around 0.5. The same holds for ETH betas, for instance w.r.t. the MVDA25 index – the CG beta is 0.568 whereas the betas w.r.t. the MVDA25 for the rest of the data sources are around 0.7. We also observe that the market betas w.r.t. the three indices are inconsistent to a degree that cannot be explained by the small differences in composition. For instance, the BTC beta for CG is 0.903 w.r.t. the CRIX but only 0.374 w.r.t. the CCi30 and 0.359 w.r.t. the MVDA25; and again, the same holds for ETH.

We were very surprised by these inconsistencies, and rather puzzled, so we decided to investigate things further and ended up writing this paper, which we intend to be used as a guide for authors of empirical studies on cryptocurrencies. First we examine the different sources of free historical crypto price data in detail. Then, finding important mistakes in some commonly-used sources, we explain how these should be corrected to avoid generating meaningless empirical results. Thereafter, we demonstrate that (even after correction) historical crypto data need handling with considerable care otherwise results will not be robust. Finally, we survey some of the literature mentioned above, finding good data practices in only a fraction of the already-published papers on cryptocurrency markets.

In the following: Section 2 provides a detailed explanation of the construction of non-traded coin prices and indices, focusing on how trading volumes distort the data provided by some coin-ranking websites; Section 3 examines the mistakes generated by Coingecko data, which (at the time of writing) feed into the CRIX index; Section 4 examines the use of non-synchronous crypto and other financial data, corrects for those exchanges which trade against tether instead of USD and identifies some interesting and recent problems emerging in Bitfinex data; Section 5 presents a volatility analysis of BTC using various models calibrated to prices from various sources, thence questioning how robust risk assessment can be in this asset class; Section 6 summarises and concludes; and some additional results are reported in appendices.

2 Volume Matters

The most common sources for cryptocurrency price data are ranking websites. A recent SEC application from Bitwise Asset Management (2019) and an article by Carter (2018) both suggest that the revenue model of such websites is largely dependent on crypto exchanges, and that consequently ranking websites are in a conflict of interest regarding their data methodologies. Retail investors decide which crypto exchange to trade on, based on the market data that ranking websites provide. Ranking websites host advertisements and referral links that 'funnel' retail investors to affiliate crypto exchanges, in return for fees. Because of the fees that certain ranking websites receive from crypto exchanges, they may be tempted to report traded volume data inflated by wash-trading, transaction fee mining, and use of non-fiat cross-rates, which may result in certain exchanges appearing more liquid than in reality. Of the data that ranking websites provide, market cap is relatively straightforward to measure, being the value of all coins in circulation – ignoring those coins held by the development team or locked in inaccessible wallets. But the daily volume traded is much more difficult to quantify for two reasons: Firstly, many exchanges deliberately inflate volume figures precisely because this boosts their ranking and attracts more traders as well as developers who are more willing to pay large fees to list their coins on these exchanges; Secondly, Coinmarketcap and Coingecko construct fiat-denominated coin indices using some sort of inferred volume from cross-trades against other crypto and different fiat rates.[7]

These artificial volume figures affect the coin values quoted by coin-ranking sites. To see how and why, we explain their methodology.[8] The dollar price index for each coin is based on the volume-weighted average of its dollar price on different crypto exchanges, so that the daily price index p_t^i for each coin i is obtained using p_t^{ij}, the price of coin i from source j at time t, and v_t^{ij}, the corresponding 24-hour volume traded from $t-1$ to t, both expressed in USD, in the formula:

$$p_t^i = \left(\sum_{j=1}^{N} v_t^{ij} \right)^{-1} \sum_{j=1}^{N} p_t^{ij} v_t^{ij} \,, \tag{1}$$

where N is the total number of price sources – for instance in the BTC/USD price index, currently N is approximately 300 for CG, 400 for CM, but only 40 for CC.

Consider the effect that each of the issues identified above may have on the individual coin indices of CC, CM and CG. Starting with volume inflation, several new exchanges have a ‘zero fee’ structure or even a ’transaction-fee mining’ structure which turns trading fees around by actually rewarding market makers for placing limit orders with the exchange’s own coin.[9] This actively encourages volume inflation because market makers earn coins through wash trading, which is still legitimate in these unregulated exchanges.[10]

Turning to volumes inferred from data on cross-rates, only Cryptocompare confines its calculation to USD-denominated crypto prices; both Coinmarketcap and Coingecko include USD prices inferred from cross-rates with other coins – including tether – and other fiat currencies. That is why they have several hundred price sources in (1). It is not clear how to infer a traded volume from

[7] See the methodologies of Coinmarketcap and Coingecko.

[8] In the following, we use the US dollar index, because it is the most commonly used rate in data analysis, but the same remarks apply to the other fiat indices provided. In fact, CC calculates price indices in multiple fiat currency denominations – e.g. a JPY-denominated BTC price index only includes BTC/JPY prices. By contrast, CG and CM denominate their price indices in USD, transforming any non-USD crypto prices into USD using FX rates. For instance if we request a JPY-denominated BTC price index from CM, they transform their USD-denominated price index into JPY using the JPY/USD FX rate.

[9] As of March 2019 the top exchanges (by monthly traded volume) that engage in transaction-fee mining were Coinbene, ZBG and FCoin. Combined monthly traded volume of transaction-fee mining exchanges represented 14% of total cryptocurrency spot traded volume in March 2019 – see Cryptocompare Exchange Review for more detail.

[10] Wash trading is the same entity (perhaps even the exchange itself) placing buy and a sell orders at the same price and volume, so that almost certainly the two orders will be executed against each other. This increases the volume without moving the price and it can be practiced at no cost (and with little risk) by any trader in zero-fee exchanges. Exchanges offering transaction-fee mining take this a step further – they calculate a trader’s fees and then reward the trader with these ‘fees’ in the form of a (supposedly) stable token set up by the exchange. Of course, until these exchanges are regulated all these questionable trading practices will be as rife as spoofing.

two cross-rates correctly, because this requires information on the size and purpose of each realised trade within the give time frame (e.g. 24-hour period). In the absence of such detailed information, ranking websites that employ cross-rates in the VWAP calculation resort to 'back-of-the-envelope' calculations to express cross-rates and their volumes in the desired base currency.[11] If the prices obtained from CM and CG behaved very similarly to the CC prices this obscurity would not be important. But using non-traded volumes inferred from cross-rates is actually an important issue – for reasons explained below when we examine the data in more detail.

Any problems with the individual coin price data will be carried over to the market-wide coin indices that employ them. The CCi30, CRIX and MVDA25 are cap-weighted indices derived from 25–50 large cap coins, typically constructed as:

$$I_t = d_t^{-1} \sum_{i=0}^{k} p_t^i \, q_s^i \,, \tag{2}$$

where: k is the number of coins included; p_t^i is the price index of coin i at time t, based on (1); q_s^i is the circulating supply of coin i at time $s \leq t$, which is typically the point when the index was last rebalanced; and the normalizing divisor d_t resets when the index composition changes.[12]

In Tables A1 and A2 of the appendices we compare the sample statistics for ordinary daily returns on BTC and ETH close prices from CG, CM, and CC, and also from Bitfinex, Coinbase, Gemini, Kraken and Poloniex for the period 1 April 2016 (1 July 2016 for ETH) – 31 March 2019. We select this period because it includes several different states of the crypto market, so that our results are not dominated by the characteristics of a particular market regime.[13] Given the variety of market states in our sample, the descriptive statistics very much depend on the sub-period examined – e.g. we observe extremely high positive returns and high volatility in the 1 April 2017 – 31 March 2018 (middle panel) which is dominated by the Q4 2017 bubble, and much lower volatility and negative returns during the bear market of 2018 (lower panel). When comparing the statistics across different data sources, the main differences are located in the skewness and excess kurtosis, which indicates that some outliers are apparent in some data sources but not in other.

Unlike other financial assets, cryptocurrency prices can also differ markedly depending on the

[11]CM first translates any non-USD prices and volumes in (1) into USD via relevant cross-rates, while ignoring the cross-rate volume. This is especially problematic for coin-to-coin cross-rates: if the prices used are all fiat-crypto rates, then the total traded volume is at least indicative of flows between crypto markets and the 'fiat world'. If crypto-to-crypto cross-rates are used, then e.g. the BTC/ETH currency pair provides an implied BTC/USD price via the ETH/USD rate; the BTC/ETH traded volume is also expressed in USD via the ETH/USD cross-rate. This method erroneously assumes that all trades in the BTC/ETH pair are executed with the purpose of exchanging the coin back to USD. Despite this serious bias, the majority of ranking websites (with the notable exception of Cryptocompare) use many non-USD prices and volumes in (1). The approach adopted by CG and Nomics is to translate the prices via cross-rates, and express all volumes in the corresponding coin. For instance, when calculating a BTC price index, traded volumes of BTC/USD, BTC/EUR, BTC/ETH etc. are all expressed in BTC. This approach does not make assumptions regarding the trading volume, although it still uses non-USD cross-rates.

[12]The CRIX and MVDA25 indices are constructed as per (2), while the CCi30 index employs a variant of (2) that weights coins by the square root of their market cap. In (2), k is 30 for the CCi30, 50 (currently) for the CRIX, and 25 for the MVDA25. As for p_t^i, it is constructed as per (1) by all market indices mentioned here. CCi30 uses CoinAPI data; CRIX uses CG price indices; MVDA25 uses CC price indices.

[13]The April 2016 – March 2019 period contains the most recent crypto market bubble of Q4 2017, the more reasonable pre-bubble bullish period until Q2 2017, the bubble's burst in January 2018, and the subsequent bearish period for the rest of 2018.

exchange. These differences are least pronounced in the most liquid market, i.e. BTC/USD, but even these prices can be significantly different across various exchanges. To justify this remark Figure A1 in the appendices exhibits the distribution of end-of-day prices over many exchanges to illustrate the extent to which synchronous prices on BTC/USD may differ. The empirical distribution is derived from 23:00:00 UTC time-stamped prices from the 12 largest exchanges (by BTC trading volume at that time) over the entire year 2018. Most percentage deviations from the Coinbase price are less than 50 basis points but some are up to 4%. Other coins have even more deviant prices, which can be more than 10% depending on the coin and exchange selected.[14]

3 Mistakes in Time Stamps

Crypto exchanges trade on a 24-7 basis so their close prices are usually measured at (or very near to) 23:59:59 UTC.[15] However, the BTC/USD daily prices from Coingecko (CG) are timestamped 00:00:00 UTC, and so are the CRIX crypto market index values because they are constructed using CG data.[16] This means that there is either a one millisecond difference, which should hardly affect the prices reported at all, or there is a whole day's difference, depending on the day that the time-stamp refers to. However, something very strange seems to have occurred on 30 January 2018 because the CG (and CRIX) prices moved out of synch with other prices since that date – and the problem still persists at the time of writing. Figure 1 illustrates the problem by plotting the CG – CC spreads for BTC and ETH (in blue) and, as a sanity check, the CM – CC spreads (in red) all relative to CC's prices.

Because of the volume issue identified above, we do not expect the spreads to be zero. However, after 30 January 2018, the CG – CC spread moved completely out of line, deviating by as much as 20%. The middle graphs of Figure 1 exhibit the same spreads but the CG prices are lagged by one day relative to the other two prices. This time, the CG spread moves into line after 30 January but is out-of-line before that date. In the lower graphs of Figure 1 we plot the CG – CC and CM – CC spreads after this mistake has been corrected. Both spreads now behave in a similar manner. We conclude that, when historical coin prices are downloaded from CG, prior to 30 January 2018 they can be used as they are; but the researcher needs to lag them starting from 30 January onwards.[17]

Now consider the market indices. In Figure 2 we compare the CRIX – CCi30 spread relative to CCi30 firstly using the CRIX data as is (upper graph), then lagging CRIX by one day (middle graph), and finally using CRIX data as is until 29 January 2018 and a 1-day lag of CRIX values starting from 30 January 2018 (lower graph). Again, the results show that lagging CRIX data after 30 January 2018 is the correct treatment. We again perform a sanity check, this time for crypto market indices. We construct the CC5, a simple price weighted index (similar in construction to

[14]Distributions of price deviations between exchanges for other coins are not reported here but are available on request.

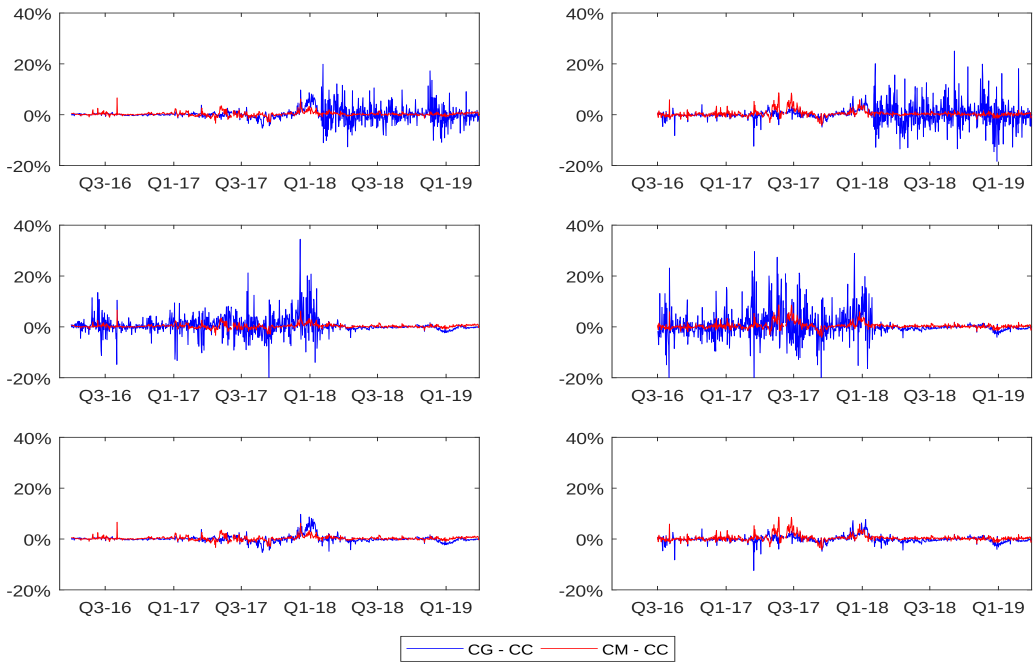
[15]Coordinated Universal Time (UTC) shares the same current time with Greenwich Mean Time (GMT), and they do not change for Daylight Saving Time (DST).

[16]We source Coingecko BTC/USD prices using its API tool – the data can be accessed at Coingecko's API. CRIX data are obtained from the CRIX website – see thecrix.de.

[17]That is, we delete the 30 January 2018 CG price and lag the remaining 31 January 2018 – 31 December 2018 price series by one day.

Figure 1: **CG – CC Spread for BTC and ETH Prices**

CG – CC and CM – CC price spreads relative to CC's daily price for BTC (left-hand graphs) and ETH (right-hand graphs), using CG price data as is (upper graphs), lagged by one day (middle graphs), and finally as is until 29 January 2018 and lagged starting from 30 January 2018 (lower graphs). The sample period is 1 April 2016 (1 July 2016 for ETH) – 31 March 2019.

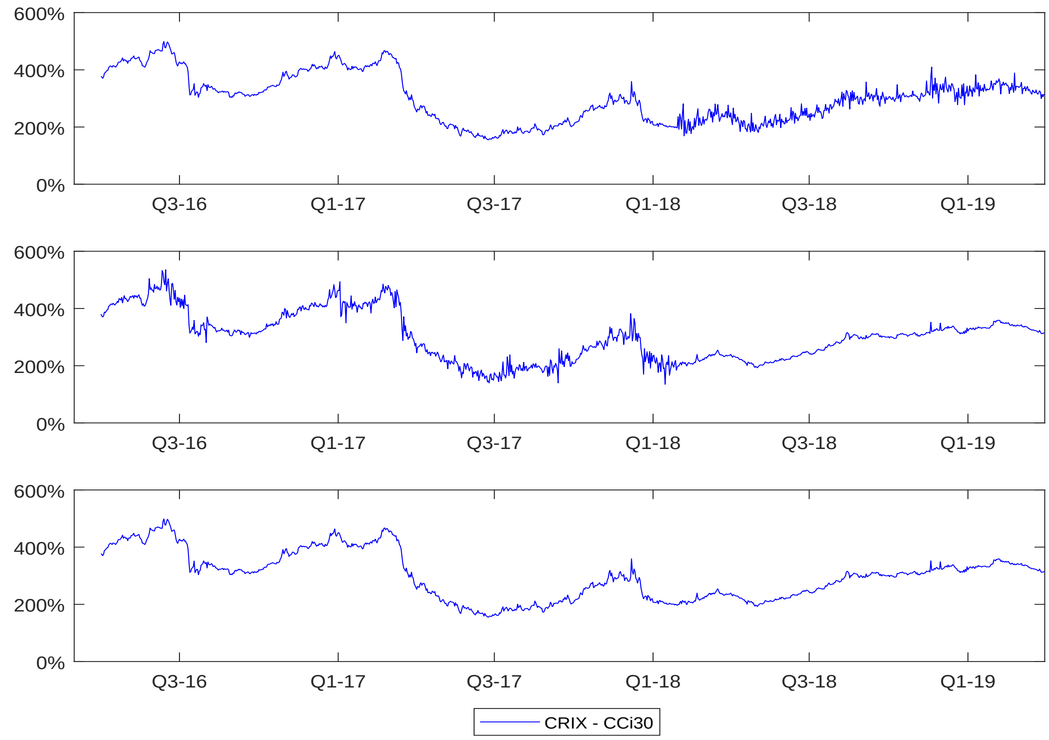


the Dow Jones stock index) that contains BTC, ETH, XRP, LTC, and DASH – the top 5 coins in market cap on 1 January 2016, using historical coin price data from CC. We find no excessive variability in the CC5 – CCi30 spread, indicating that the CCi30 data are not at fault. However, it is clear the CRIX data have absorbed the mistake in the CG data as described above.[18] As of June 2018 four published papers and other discussion papers employ Coingecko as a data source, and a further eleven papers employ CRIX index data.

[18]It would be good to use ready-made crypto market indices with a more sophisticated construction methodology such as the MVDA25 or the Bletchley 10 index – see www.bletchleyindexes.com. However, this is not possible. The MVDA25 values refer to 17:00:00 UTC, and the Bletchley 10 index does not provide historical data as far back as early 2016.

Figure 2: **CCi30 − CRIX Spread**

CRIX – CCi30 spread relative to CCi30, using CRIX daily data as is (upper graph), lagged by one day (middle graph), and as is until 29 January 2018 and lagged starting from 30 January 2018 (lower graph). The sample period is 1 April 2016 – 31 March 2019.



4 Other Data Issues

In this Section we begin by examining the use of non-synchronous crypto and other financial data, and then we discuss the deviations between traded prices which sometimes behave quite strangely.

We now examine the use of non-synchronous crypto and other financial data, which can be problematic due to the high volatility of crypto prices. But even when using synchronous traded data, prior to 2018 there were significant price deviations across crypto exchanges; these deviations have become very small since January 2018, with the exceptions of tether-denominated crypto prices which still exhibit significant deviations.

If crypto prices are used in multivariate analysis with other types of financial assets it is necessary to access data on an intra-day basis in order to obtain synchronous prices across all assets used.[19] A few hours difference in the time of price measurement could be particularly problematic in crypto markets because of the very high volatility of coin prices. That is, data are used in a correlation or other multivariate analysis where the price of a coin is for instance registered 3 or 4 hours after the close price of a New York-based stock index. Typically the time lag can range from 3 to 18 hours.[20] Even in BTC, which is the least volatile of the commonly traded cryptocurrencies, the

[19] For instance, the close prices on the New York Stock Exchange (NYSE) are only available Monday through Friday at 16:00 Eastern Time (UTC-05:00 or UTC-04:00 depending on Daylight Saving Time). Another advantage of CC is that it allows high-frequency data to be downloaded.

[20] For instance, the S&P 500 close time is 16:00 Eastern Time (UTC-05:00 or UTC-04:00 in DST), so crypto prices are reported 3 or 4 hours later; the Euro Stoxx 50 close time is 18:00 Central European Time (UTC+01:00 or

average intra-day price range over the period 1 January 2016 – 31 December 2018 is 5.85% with a positive skew and heavy-tails, as shown in Figure A2 of the appendices; and for ETH the average intra-day range in the same sample period is 10.5%.[21]

It is important to use concurrent data for papers studying portfolio diversification and hedging, but also for any other paper analysing crypto and other financial data in combination. Borri and Shakhnov (2018), Karalevicius (2018), Baur and Dimpfl (2019), Borri (2019) and Urquhart and Zhang (2019) acknowledge this issue and treat it accordingly.[22] It is also particularly important to use synchronous data for studies of arbitrage. Some papers on exchange arbitrage are careful to use synchronous high-frequency data. For instance, Lintilhac and Tourin (2017) use 5-hour frequency traded data on BTC-e, Bitstamp and itBit with a sample period of 4 January 2014 to 3 June 2016 to demonstrate profitable arbitrage opportunities after accounting for bid/ask slippage, transactions costs and market impact costs. Similarly, Makarov and Schoar (2018) use tick level trading data from the 15 largest and most liquid exchanges between 1 January 2017 – 28 February 2018, and find large arbitrage opportunities across different exchanges that often persist for several days. Also, when examining cryptocurrency futures and their role in price discovery, synchronous data must be used – see e.g. (Alexander and Heck, 2019; Baur and Dimpfl, 2019; Choi et al., 2019).[23]

Significant price deviations between crypto exchange prices allowed arbitrageurs to extract profits, but since the beginning of 2018 these price deviations have become much smaller. To see this Figure 3 depicts the price spread between the Coinbase BTC price and four other long-established centralised exchanges, always relative to the Coinbase price. The spreads are large and variable until the second quarter of 2018, after which they become almost negligible.[24] However, in the last

UTC+02:00 in UTC), so crypto prices are reported 7 or 8 hours later; the Nikkei 225 close time is 15:00 UTC+09:00 (no DST), so crypto prices are reported 18 hours later; the CSI 300 close time is 15:00 UTC+08:00 (no DST), so crypto prices are reported 17 hours later. Crypto markets operate on a 24-7 basis with almost no exceptions, so a common convention in historical price data is to report each day's 'close' price at or very near 23:59:59 UTC.

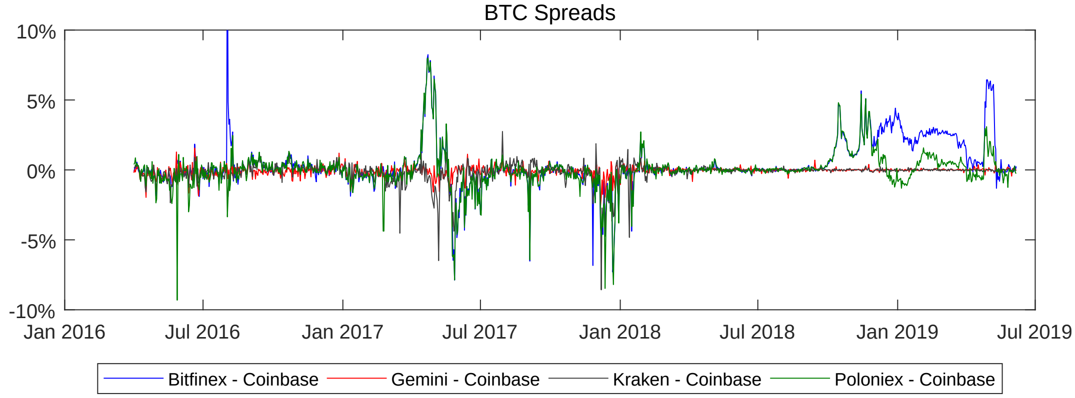
[21]Of course, those cryptocurrencies that are hardly traded at all have illiquid markets and so their prices do not move much, but we are not interested in those.

[22]Borri and Shakhnov (2018) examine the efficiency of bitcoin markets. They compute end-of-day bitcoin prices corresponding to 16:00 UTC and drop observations corresponding to weekends and additional non-business days, to match bitcoin prices with fiat currency exchange rates (obtained from Thompson Reuters) whose daily close occurs at 16:00 UTC; and the same treatment is applied in Borri (2019). Karalevicius (2018) examines whether investor sentiment is a good predictor for bitcoin's returns. He webscrapes news articles from Coindesk, Cointelegraph and other crypto news websites, and takes care to convert all dates and times to UTC+00:00 in accordance with the price data he collects from Bitstamp so that sentiment analysis is performed on the correct subsample of articles for each day. Baur and Dimpfl (2019) use 15-minute frequency for their bitcoin spot price data, in order to align them exactly on the bitcoin futures trading hours. Urquhart and Zhang (2019) examine bitcoin's role as a hedge or diversifier against fiat currency exchange rates. Since foreign exchange trades on a 24-hour basis except weekends, and bitcoin trades on a 24-7 basis, they choose to filter out bitcoin prices during periods when the currency markets are closed.

[23]Other papers use synchronous spot data to study crypto market efficiency: Bariviera (2017), Vidal-Tomás and Ibañez (2018) and Zargar and Kumar (2019) use data from Bitstamp; Pieters and Vivanco (2017) and Sensoy (2019) use data from multiple crypto exchanges, as can be obtained from Bitcoincharts; and Borri and Shakhnov (2018) consider prices from over 100 exchanges listed on Cryptocompare; Brauneis et al. (2018) access the REST API of Bitfinex, Bitstamp and Coinbase, requesting both high-frequency trades and limit order book data.

[24]All significant deviations of these spreads from zero can be tied with extreme events in the crypto market. For instance, on 28 May 2016 BTC traded at a significant discount on Poloniex just when the enormously-hyped DAO token began trading on Poloniex. On 2 August 2016 BTC the Bitfinex – Coinbase BTC spread spiked because Bitfinex halted trading due to a hack and at the same time BTC prices crashed on most other exchanges. In April 2017 Bitfinex and Poloniex BTC prices decoupled from other exchanges due to issues faced by Bitfinex and Tether with their banking partners. And in the fourth quarter of 2017 Coinbase BTC prices were higher than Bitfinex,

Figure 3: **Bitfinex, Gemini, Kraken and Poloniex BTC Spreads against Coinbase**
Bitfinex – Coinbase, Gemini – Coinbase, Kraken – Coinbase and Poloniex – Coinbase BTC price spreads, all expressed relative to Coinbase's BTC price. The data frequency is daily and the sample period is 1 April 2016 – 26 June 2019.



quarter of the sample period when the spreads on Bitfinex and Poloniex diverge compared with the spreads on Gemini and Kraken, because the prices on Bitfinex and Poloniex are expressed in tether (USDT) not USD. Tether is a stablecoin and is supposed to be exchangeable 1:1 with USD, but in fact it has been known to deviate significantly, offering quite a lot of arbitrage opportunities. In fact, on 15 January 2015 Bitfinex announced that it accepts USDT deposits and credits them as USD to users' trading accounts at a 1:1 rate. The tether-dollar parity was effectively ended by Bitfinex on 27 November 2018 when it introduced a USDT/USD cross-rate, and even more so on 21 December 2018 when they introduced margin trading on that pair with up to 3.3x leverage. But until 11 March 2019 when separate BTC/USDT and ETH/USDT pairs were introduced by Bitfinex, BTC and ETH Bitfinex prices were effectively denominated in tether.

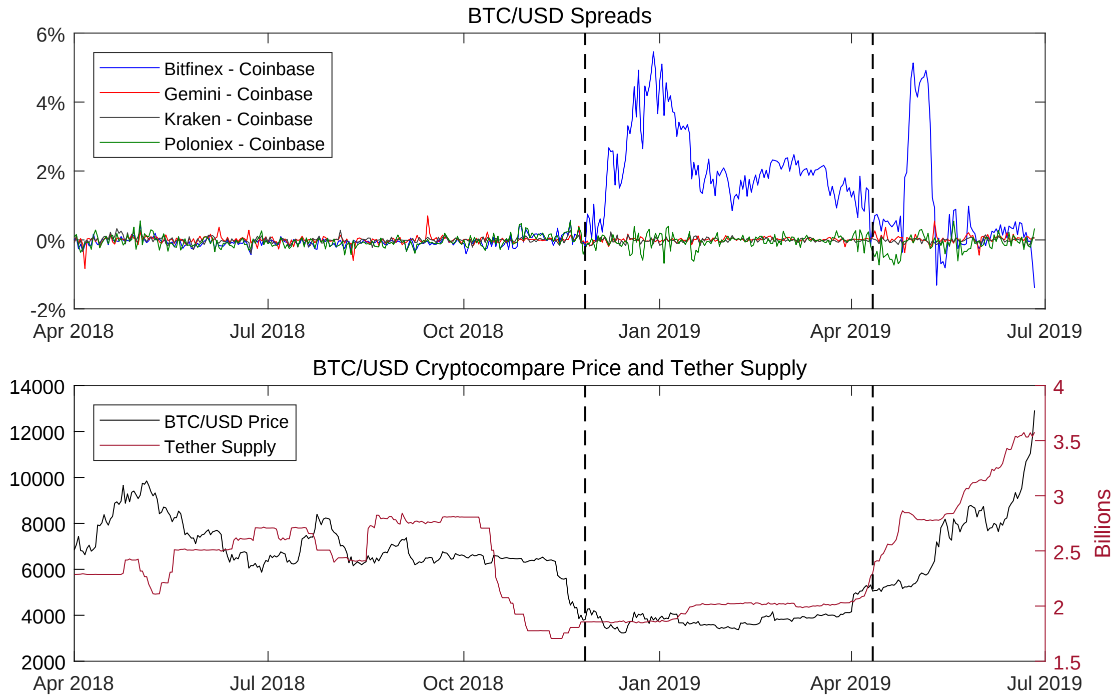
In the upper graph of Figure 4 we express the Bitfinex and Poloniex BTC prices in USD using the USDT/USD cross-rate available on Kraken.[25] The Poloniex prices (in green) move into line with the rest after adjustment, but the Bitfinex prices (in blue) do not, and even using the Bitfinex USDT/USD price to adjust the Bitfinex BTC price does not bring it back in line with the others. Interestingly, the decoupling of the Bitfinex BTC price (even after tether adjustment) coincides exactly with a significant decline in BTC/USD price levels during the fourth quarter of 2018 as shown in Figure 4. This event occurs very close to the introduction of the USDT/USD pair by Bitfinex on 27 November 2018, denoted by the first dotted line in the upper graph of Figure 4; the second dotted line denotes the introduction of margin trading for the BTC/USDT and ETH/USDT Bitfinex pairs (11 April 2019). The Bitfinex price subsequently converges towards the Coinbase price, making the Bitfinex – Coinbase BTC price spread in the upper graph very small. A second Bitfinex price decoupling occurs in mid-April 2019 (upper graph), and coincides very closely with the legal issues face by Bitfinex. Most alarming of all, the lower graph of Figure 4 depicts an immense increase in the supply of tether by more than 1.5 billion tokens between mid-March and late June 2019. This increase in tether's supply is accompanied by an impressive surge in BTC/USD

Kraken and Poloniex prices, possibly because many of its new users traded through its retail platform.

[25]We use the Kraken USDT/USD price because it has much higher (real) volume compared with other crypto exchanges and the longest historical period available.

prices across all exchanges – see Alexander and Dakos (2019) for more detail, and Griffin and Shams (2018) for empirical evidence of tether's role in the bitcoin price bubble of 2017.

Figure 4: **BTC Tether-Adjusted Price Spreads, BTC Price Level and USDT Supply**
Upper graph: Bitfinex – Coinbase, Gemini – Coinbase, Kraken – Coinbase and Poloniex – Coinbase BTC price spreads, all expressed relative to Coinbase's BTC price. Bitfinex and Poloniex prices are expressed in USD via the Kraken USDT/USD cross-rate. **Lower graph:** BTC price index from CC. The data frequency is daily, and the sample period is 1 April 2018 – 26 June 2019. Vertical dotted lines denote dates of interest.



5 Challenges in Risk Assessment

Here we demonstrate that the choice of data source has an enormous influence on the calibration of statistical volatility models for cryptocurrencies, using BTC and ETH for our empirical results. Numerous papers explore the best specification for generalised autoregressive conditional heteroscedasticity (GARCH) models introduced by Bollerslev (1986) on different types of financial data, see Engle et al. (2008) for a useful survey. Ten recent papers have explored this topic using returns on BTC and other cryptos. Bouoiyour and Selmi (2016) examine several specifications for BTC returns and find that the optimal specification is the component with multiple threshold (CMT)-GARCH. Katsiampa (2017) claims that asymmetric component GARCH fits BTC returns better than many other models. Corbet et al. (2018) use symmetric GARCH and Vidal-Tomás and Ibañez (2018) use a component GARCH, both with un-specified distributions for innovations. By contrast, Bouri et al. (2017) use a symmetric model with innovations that follow a generalised error distribution (GED), and Al-Khazali et al. (2018) attempt several specifications, finding that the optimal model is the exponential GARCH with normally distributed residuals. Dyhrberg (2016) uses a symmetric and an exponential normal GARCH on BTC returns. Similarly, Baur et al. (2018) use a normal exponential GARCH model, and note that using a Student-t distribution does not improve their results. They also use a normal GJR-GARCH and find that estimates violate the parameter constraints for a finite positive steady-state volatility, so the variance process is integrated.[26]

All of the above papers calibrate models using maximum likelihood estimation (MLE). However this optimizer lacks robustness: when we generate observations from a GARCH model with fixed parameters and use these observations to re-estimate the same GARCH specification, MLE cannot identify the same parameters. Following Asai (2006), Virbickaite et al. (2015) and Ardia (2008) it is now recognised that Markov chain Monte Carlo (MCMC) is much more robust than MLE for calibrating univariate GARCH data generation processes. Using MCMC, Ardia et al. (2019) motivate the use of Markov switching GARCH models for BTC data, and find that a two-regime asymmetric (GJR) skewed Student-t GARCH model provides the best in-sample fit. Also using MCMC, Caporale and Zekokh (2019) extend these results to ether, ripple and litecoin, using Value-at-Risk and Expected Shortfall backtesting and the model confidence set procedure to select the best model specification.

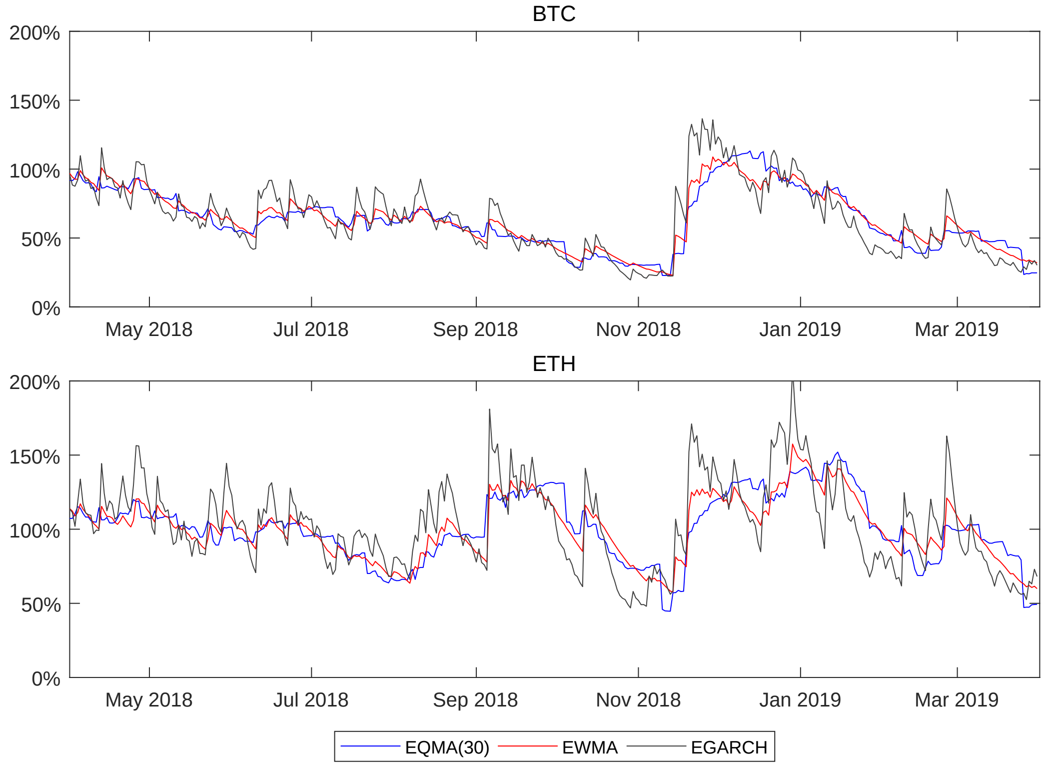
Before we put these risk models for cryptocurrencies under the microscope, we ask how important it is to get the right risk model – e.g. perhaps a moving average estimation is enough. Suppose we take a 30-day equally-weighted moving average (EQMA) or a standard RiskMetrics exponentially weighted moving average (EWMA) with smoothing constant 0.94 – would we obtain similar results to those generated by a GARCH model? Let us compare the different model's in-sample volatility estimates for daily returns on BTC and ETH prices from CC. In Figure 5 we plot the 30-day EQMA, the EWMA with $\lambda = 0.94$ and – after much experimentation with several GARCH specifications – the exponential GARCH model of Nelson (1991) with Student-t distributed innovations.[27] Bitcoin's volatility is lower than other cryptos; ether's volatility is higher, yet it is the second least-volatile cryptocurrency. The spreads in volatility estimates between different methods can be as high as 400% for BTC and 100% for ETH, and they maximize during periods of high volatility. For instance, on 21 November 2018 BTC volatility is 73% when calculated with an EQMA(30) and 133% when calculated with a single-state Student-t EGARCH. This has an enormous impact on Sharpe ratio calculation that most funds report on a regular basis: PwC reports an average quant crypto hedge fund return of a little more than 10% p.a. in 2018; assuming that a funds return on 21 Novemeber 2018 was 10% p.a. and with current borrowing rates so low, this would translate into an ex-post SR of either 0.14 using the EQMA(30) or 0.07 using the single-state Student-t EGARCH. By the same token, such huge differences in volatility, depending on risk model choice, would have a significant effect on the capital provisions necessary for large institutional investors. In view of these findings we conclude that the choice of risk model is very important indeed – a risk manager should deploy a well-defined GARCH model which should be calibrated using MCMC rather than MLE.

Now we investigate how our choice of data source affects the GARCH model calibration. For this, we use daily returns on BTC prices obtained from various sources. We find similar results to Baur et al. (2018) for non-switching GARCH specifications, i.e. the estimated variance process is

[26] Also Borri (2019) examines conditional tail-risk in the markets for bitcoin, ether, ripple and litecoin using prices obtained from the Bitfinex and Bitstamp exchanges. Matkovskyy (2019) compares the EUR, USD, and GBP centralized and decentralized BTC crypto exchanges in terms of return volatility and interdependency.

[27] We were only able to calibrate this model by imposing a reasonable value for the degrees of freedom parameter based on sample statistics moments.

Figure 5: **Cryptocompare BTC and ETH In-sample Volatility**
We depict the 30-day EQMA, EWMA and Student-t EGARCH in-sample volatility estimates based on the daily returns of BTC (upper graph) and ETH (lower graph) prices from CC. The sample period is 1 April 2018 – 31 March 2019.



almost always integrated and most models are clearly misspecified.[28] Therefore, we explore Markov switching specifications. Following Ardia et al. (2019) we use the two-state Markov switching GJR-GARCH model and estimate its parameters using MCMC.[29] We filter the log returns through an AR(1) conditional mean and then apply the two-state Markov switching skewed Student-t GJR-GARCH model:

$$\sigma_{it}^2 = \omega_i + \left(\alpha_i + \gamma_i I_{\{\varepsilon_{t-1}<0\}}\right)\varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2 \text{ where } \varepsilon_t \sim t_{\eta_i,\xi_i} \text{ and } i = 1, 2.$$

We approximate the unconditional steady-state volatility of each regime as:

$$UV_i = \sqrt{\frac{365\,\omega_i}{1 - \alpha_i - 0.5\gamma_i - \beta_i}}.$$

The β parameter controls the persistence of price shocks in the volatility, and volatility may react differently to positive and negative price shocks depending on the values of α and γ. The degrees

[28]Detailed results are available on request.
[29]Ardia et al. (2019) take log returns on BTC Bitstamp prices, ranging from 19 August 2011 until 2 March 2018. They find that out of several model specifications the two-state skewed Student-t GJR-GARCH provides the best in-sample fit based on the Deviance Information Criterion (DIC). We can successfully replicate their results using the MSGARCH R package documented in Ardia et al. (2019). We thank Dr. David Ardia and Keven Bluteau for their valuable help in adapting the MSGARCH R package for the leverage effect we need for crypto returns.

of freedom (η) and asymmetry (ξ) parameters for the skewed Student-t distribution are also state-dependent. The constant state transition probabilities control the latent state variable s_t; these are summarised in the transition probability matrix:

$$\mathbf{\Pi} = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix},$$

where each element $p_{ij} = P\left(s_t = j | s_{t-1} = i\right)$ denotes the constant probability of transitioning from state i at time $t-1$ to state j at time t.

To understand the 'data source effect' on the GARCH model parameters, we use BTC prices from CG, CM, CC, Bitstamp and Kraken between 1 September 2013 – 31 March 2019.[30] We produce the posterior parameter distribution using 10,000 burn-in draws and build a sample of size 2,000 with the next 10,000 draws keeping only every 5^{th} draw, setting our pseudo-random number generator seed to 1 for each data source. Table 2 reports the posterior distribution median and (in parentheses) the 25^{th} and 75^{th} percentiles for each parameter in the Markov switching skewed Student-t GJR-GARCH. The parameter estimates are highly sensitive to the data source choice. For instance, in state 1 which is the low volatility regime (upper panel of Table 2) the CG ω_1 median falls outside CC's $(25^{th}, 75^{th})$ percentile interval, even though the two medians are quite close; and similarly for γ_1 and p_{11}. Even worse, in state 2 which is the high volatility regime, the median estimate for the leverage parameter γ_2 does not even maintain the same sign for all sources, and the corresponding $(25^{th}, 75^{th})$ percentile intervals indicate a lot of dispersion in the parameter distributions; the estimates of β in state 2 based on the CG and Bitstamp data are both well below the normal values for a GARCH persistence parameter; and the steady-state volatility estimates vary considerably across data sources. In fact all the calibrated parameters differ, in some cases very considerably, across different data sources.

Given the large discrepancies in parameter estimates across data sources, we now ask whether the optimal model choice also varies when we use different data. We compare the optimal GARCH model specification for BTC returns from CG, CM, CC, Bitstamp and Kraken, as indicated by the Deviance Information Criterion (DIC) and the Bayesian Predictive Information Criterion (IC). Table A3 of the appendices reports the DIC and IC for the following model specifications estimated using MCMC: two-state Markov switching GARCH, GJR-GARCH or EGARCH with innovations distributed according to a (symmetric or skewed) normal or Student-t distribution. For each data source we highlight (in blue) the minimum DIC and IC values which indicate the optimal model choice from those examined. The DIC and IC reported in Table A3 indicate that the optimal models are almost exclusively the two-state Markov switching symmetric GARCH and EGARCH with innovations that follow a skewed Student-t distributions. These results are different compared with Ardia et al. (2019), but we do not employ the same historical period.

Because the optimal choice of model can be highly sensitive to the sample period, we also examine the symmetric GARCH and EGARCH specifications as defined by Bollerslev (1986) and Nelson

[30] We do not use ETH data since they are only available from mid-2015 onwards, and the sample length is not sufficient for a Markov switching model calibration, even with the MCMC method.

Table 2: **Markov Switching GJR-GARCH on Non-Traded and Traded BTC Data**
Markov switching skewed Student-t GJR-GARCH parameter medians and (in parentheses) the 25^{th} and 75^{th} percentiles, based on BTC daily returns from CG, CM, CC, Bitstamp and Kraken. The sample period is 1 September 2013 – 31 March 2019. Parameters of interest are highlighted in blue and red.

	CG	**CM**	**CC**	**Bitstamp**	**Kraken**
			State 1		
ω_1	0.145 (0.116, 0.185)	0.007 (0.006, 0.007)	0.197 (0.166, 4.932)	0.036 (0.031, 0.042)	0.011 (0.009, 0.014)
α_1	0.147 (0.126, 0.168)	0.052 (0.046, 0.058)	0.156 (0.051, 0.181)	0.070 (0.051, 0.093)	0.100 (0.085, 0.117)
γ_1	-0.067 (-0.091, -0.035)	-0.019 (-0.034, -0.005)	-0.097 (-0.144, 0.246)	-0.006 (-0.033, 0.020)	-0.074 (-0.089, -0.061)
β_1	0.865 (0.846, 0.878)	0.953 (0.947, 0.958)	0.858 (0.662, 0.877)	0.922 (0.910, 0.934)	0.917 (0.908, 0.926)
η_1	2.948 (2.846, 3.093)	2.396 (2.362, 2.441)	2.873 (2.709, 4.653)	2.832 (2.699, 3.001)	3.057 (2.950, 3.202)
ξ_1	1.013 (0.989, 1.033)	1.046 (1.026, 1.068)	0.998 (0.902, 1.031)	1.028 (1.008, 1.046)	1.100 (1.082, 1.124)
p_{11}	0.944 (0.934, 0.952)	0.957 (0.950, 0.964)	0.976 (0.972, 0.980)	0.975 (0.970, 0.979)	0.862 (0.846, 0.879)
UV_1	49.6%	21.5%	45.7%	36.3%	14.2%
			State 2		
ω_2	30.442 (27.905, 32.819)	2.377 (2.129, 2.645)	4.599 (0.168, 5.435)	24.312 (20.950, 28.351)	1.828 (1.675, 2.004)
α_2	0.003 (0.003, 0.004)	0.240 (0.222, 0.258)	0.067 (0.054, 0.185)	0.231 (0.197, 0.265)	0.022 (0.019, 0.025)
γ_2	0.297 (0.230, 0.364)	-0.055 (-0.082, -0.031)	0.204 (-0.144, 0.262)	0.311 (0.185, 0.424)	0.168 (0.140, 0.202)
β_2	0.009 (0.004, 0.028)	0.767 (0.754, 0.779)	0.682 (0.657, 0.880)	0.308 (0.275, 0.348)	0.880 (0.862, 0.896)
η_2	51.933 (29.060, 70.020)	4.172 (3.895, 4.542)	4.165 (2.684, 4.835)	3.131 (2.860, 3.466)	65.485 (61.535, 69.464)
ξ_2	0.710 (0.661, 0.775)	0.865 (0.837, 0.893)	0.937 (0.880, 1.022)	0.854 (0.813, 0.891)	0.906 (0.851, 0.960)
p_{22}	0.772 (0.730, 0.815)	0.950 (0.944, 0.956)	0.974 (0.969, 0.979)	0.951 (0.936, 0.961)	0.599 (0.569, 0.634)
UV_2	115%	204%	106%	170%	216%

(1991) respectively, reporting results in the appendices (and results for other GARCH models are available on request). Specifically, the MCMC parameter medians (and percentiles) of the skewed Student-t symmetric GARCH and EGARCH models are reported in Tables A4 and A5 of the appendices. Again we note significant discrepancies in all parameters across data sources. For instance, the skewed Student-t GARCH appears to be the optimal model for CG and Bitstamp according the corresponding DIC and IC values in Table A3. However, when we examine the parameter estimates in Table A4 every parameter median for CG is outside the (25^{th}, 75^{th}) percentile interval of the same parameter for Bitstamp, except for the distribution parameters η and ξ, and the state 2 transition probability p_{22}; and the same holds vice versa. Similarly, the skewed Student-t EGARCH is the optimal model for CC and Kraken data. And yet, when we examine the EGARCH parameter medians and percentile intervals of CC and Kraken displayed in Table A5 the parameter medians of CC are outside the percentiles' interval of Kraken, and vice versa, for almost all parameters.

6 Summary and Conclusions

In the light of all our findings and discussions, we critically surveyed the last two and a half years of papers published in finance and economics journals, and some of the more recent SSRN papers. We summarise those papers which use the wrong kind of data without citing any of them explicitly. As explained in Section 1, by 'wrong' we mean any of the following: data from questionable sources; non-concurrent time-series data in multivariate analysis; non-traded prices in portfolio optimisation, efficiency studies, trading strategy development or hedging analysis.

In total, we examined 152 published and SSRN discussion papers. Of these, 38 use data from questionable sources; 38 incorrectly employ non-traded prices; and 36 use non-synchronous data in a multivariate analyses across different asset classes. Approximately half the relevant published papers in our literature search (67 from a total of 124) are published in *Economics Letters*, *Finance Research Letters* and *Research in International Business and Finance*; 39 of these use 'wrong' data. Having said this, out of the 152 published and discussion papers in our search, 25 papers *do* use proper traded prices when examining the topics mentioned above.[31] Hopefully by the time this article is published, there will be many more.[32] As for data synchroneity, 5 papers in our sample *do* use concurrent data.[33]

When examining crypto market efficiency, portfolio optimization, hedging and trading applications it is very important to use traded data from crypto exchanges, not data from coin-ranking sites. But some exchanges do not trade fiat currencies and instead use 'stablecoins' such as tether. So their prices need adjustment in order to be comparable with fiat prices. Even then, significant inconsistencies can occur between traded prices on different exchanges. Researchers and traders need to be aware of such inconsistencies both in order to avoid using 'compromised' data, and because these cases may be excellent topics for further research. The BTC/USD price premium of Bitfinex and its potential implications form one such case that we discuss in brief.

It is also important to use a well-conditioned GARCH model for risk assessment of cryptocurrencies. Following Ardia et al. (2019) we support the use of Student-t Markov-switching GARCH

[31] These are the following: Bariviera (2017), Lintilhac and Tourin (2017), Pieters and Vivanco (2017), Urquhart (2017), Baur et al. (2018), Benedetti (2018), Borri and Shakhnov (2018), Brauneis et al. (2018), Dyhrberg et al. (2018), Griffin and Shams (2018), Koutmos (2018), Makarov and Schoar (2018), Vidal-Tomás and Ibañez (2018), Baig et al. (2019), Baur and Dimpfl (2019), Borri (2019), Borri and Shakhnov (2019), Bouri et al. (2019), Choi et al. (2019), Eross et al. (2019), Hu et al. (2019), Matkovskyy (2019), Mbanga (2019), Sensoy (2019), Urquhart and Zhang (2019), and finally Zargar and Kumar (2019).

[32] These numbers are valid as of the time of writing, and to the best of our knowledge.

[33] These are: Borri and Shakhnov (2018), Borri (2019), Karalevicius (2018), Baur and Dimpfl (2019) and Urquhart and Zhang (2019).

models with two states, at least for BTC returns. However, we find that the choice between a symmetric or asymmetric volatility model very much depends on the data source as well as the sample period, and so do the model parameters themselves. In practice the risk manager must be very careful to calibrate this model on data from the crypto exchange that they are actually trading on. Otherwise, the parameter estimates are likely to be quite incorrect. For instance, if a trader uses CC or Bitstamp prices for investments made on Kraken, they will get very different results. And when a non-traded coin price can be used for risk assessment we recommend using the Cryptocompare data because other coin-ranking sites base their quotes on unreliable volume data.

In finance we sometimes tend to worry very little about the quality of our data; after all, Bloomberg is Bloomberg. However, the choice of data source is of particular importance in cryptocurrency research. We have tried to provide a framework for best data practice which, having surveyed the recent literature, is clearly necessary in this area. There are numerous different sources of freely-downloadable data to choose from but some are clearly better than others and we have given several reasons why. In addition to volume-inflation, which can distort some individual coin data from coin-ranking sites, the Coingecko data contains significant errors which feed into the CRIX market index. We have shown how these errors can be corrected, and that it is necessary to do this before any meaningful empirical analysis can be performed.

References

Al-Khazali, O., E. Bouri, and D. Roubaud (2018). The impact of positive and negative macroeconomic news surprises: Gold versus bitcoin. *Economics Bulletin 38*(1), 373–382.

Alexander, C. and M. Dakos (2019). Bitcoin: welcome to Bitfinex's second Tether bubble. Medium.

Alexander, C. and D. Heck (2019). Price discovery and efficiency in bitcoin markets. SSRN Discussion Paper 3383147.

Ardia, D. (2008). *Financial risk management with bayesian estimation of GARCH models*. Springer-Verlag Berlin Heidelberg.

Ardia, D., K. Bluteau, K. Boudt, L. Catania, and D.-A. Trottier (2019). Markov-Switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software*. Forthcoming.

Ardia, D., K. Bluteau, and M. Rüede (2019). Regime changes in bitcoin garch volatility dynamics. *Finance Research Letters 29*, 266–271.

Asai, M. (2006). Comparison of MCMC methods for estimating GARCH models. *Journal of the Japan Statistical Society 36*(2), 199–212.

Baig, A., B. Blau, and N. Sabah (2019). Price clustering and sentiment in bitcoin. *Finance Research Letters 29*, 111–116.

Bariviera, A. (2017). The inefficiency of bitcoin revisited: A dynamic approach. *Economics Letters 161*, 1–4.

Baur, D. and T. Dimpfl (2019). Price discovery in bitcoin spot or futures? *Journal of Futures Markets 39*(7), 803–817.

Baur, D., T. Dimpfl, and K. Kuck (2018). Bitcoin, gold and the US dollar: A replication and extension. *Finance Research Letters 25*, 103–110.

Baur, D., K. Hong, and A. Lee (2018). Bitcoin: Medium of exchange or speculative assets? *Journal of International Financial Markets, Institutions and Money 54*, 177–189.

Benedetti, H. (2018). The economics of digital token cross-listings. SSRN Discussion Paper 3267392.

Bitwise Asset Management (2019). Presentation to the U.S. Securities and Exchange Commission. Available at: https://bit.ly/2TSN9oi.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Borri, N. (2019). Conditional tail-risk in cryptocurrency markets. *Journal of Empirical Finance 50*, 1–19.

Borri, N. and K. Shakhnov (2018). Cryptomarket discounts. SSRN Discussion Paper 3124394.

Borri, N. and K. Shakhnov (2019). Regulation spillovers across cryptocurrency markets. SSRN Discussion Paper 3343696.

Bouoiyour, J. and R. Selmi (2016). Bitcoin: A beginning of a new phase? *Economics Bulletin 36*(3), 1430–1440.

Bouri, E., L. Gil-Alana, R. Gupta, and D. Roubaud (2019). Modelling long memory volatility in the bitcoin market: Evidence of persistence and structural breaks. *International Journal of Finance and Economics 24*(1), 412–426.

Bouri, E., P. Molnar, G. Azzi, D. Roubaud, and L. Hagfors (2017). On the hedge and safe haven properties of bitcoin: Is it really more than a diversifier? *Finance Research Letters 20*, 192–198.

Brauneis, A., R. Mestel, R. Riordan, and E. Theissen (2018). A high-frequency analysis of bitcoin markets. SSRN Discussion Paper 3249477.

Caporale, G. and T. Zekokh (2019). Modelling volatility of cryptocurrencies using markov-switching garch models. *Research in International Business and Finance 48*, 143–155.

Carter, N. (2018). A glimpse into the dark underbelly of cryptocurrency markets. Medium.

Choi, J., H. Park, and S. Sohn (2019). Price discovery in bitcoin futures: Evidence from BitMEX. SSRN Discussion Paper 3353583.

Corbet, S., B. Lucey, M. Peat, and S. Vigne (2018). Bitcoin futures: What use are they? *Economics Letters 172*, 23–27.

Daian, P., S. Goldfeder, T. Kell, Y. Li, X. Zhao, I. Bentov, L. Breidenbach, and A. Juels (2019). Flash Boys 2.0: Frontrunning, transaction reordering, and consensus instability in decentralized exchanges. *CoRR abs/1904.05234*.

Dyhrberg, A. (2016). Bitcoin, gold and the dollar - A GARCH volatility analysis. *Finance Research Letters 16*, 85–92.

Dyhrberg, A., S. Foley, and J. Svec (2018). How investible is bitcoin? Analyzing the liquidity and transaction costs of bitcoin markets. *Economics Letters 171*, 140–143.

Engle, R., S. Focardi, and F. Fabozzi (2008). ARCH/GARCH models in applied financial econometrics. *Handbook of Finance, John Wiley and Sons*. Edited by Frank J. Fabozzi.

Eross, A., F. McGroarty, A. Urquhart, and S. Wolfe (2019). The intraday dynamics of bitcoin. *Research in International Business and Finance 49*, 71–81.

Griffin, J. M. and A. Shams (2018). Is bitcoin really un-tethered? SSRN Discussion Paper 3195066.

Hu, B., T. McInish, J. Miller, and L. Zeng (2019). Intraday price behavior of cryptocurrencies. *Finance Research Letters 28*, 337–342.

Karalevicius, V. (2018). Using sentiment analysis to predict interday bitcoin price movements. *Journal of Risk Finance 19*(1), 56–75.

Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of GARCH models. *Economics Letters 158*, 3–6.

Koutmos, D. (2018). Liquidity uncertainty and bitcoin's market microstructure. *Economics Letters 172*, 97–101.

Lintilhac, P. and A. Tourin (2017). Model-based pairs trading in the bitcoin markets. *Quantitative Finance 17*(5), 703–716.

Makarov, I. and A. Schoar (2018, Apr). Trading and arbitrage in cryptocurrency markets. SSRN Discussion Paper 3171204.

Matkovskyy, R. (2019). Centralized and decentralized bitcoin markets: Euro vs USD vs GBP. *Quarterly Review of Economics and Finance 71*, 270–279.

Mbanga, C. (2019). The day-of-the-week pattern of price clustering in Bitcoin. *Applied Economics Letters 26*(10), 807–811.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica 59*(2), 347–370.

Pieters, G. and S. Vivanco (2017). Financial regulations and price inconsistencies across bitcoin markets. *Information Economics and Policy 39*, 1–14.

Sensoy, A. (2019). The inefficiency of bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters 28*, 68–73.

Trimborn, S. and W. K. Härdle (2018). CRIX an index for cryptocurrencies. *Journal of Empirical Finance 49*, 107–122.

Urquhart, A. (2017). Price clustering in bitcoin. *Economics Letters 159*, 145–148.

Urquhart, A. and H. Zhang (2019). Is bitcoin a hedge or safe haven for currencies? An intraday analysis. *International Review of Financial Analysis 63*, 49–57.

Vidal-Tomás, D. and A. Ibañez (2018). Semi-strong efficiency of bitcoin. *Finance Research Letters 27*, 259–265.

Virbickaite, A., M. Ausn, and P. Galeano (2015). Bayesian inference methods for univariate and multivariate GARCH models: A survey. *Journal of Economic Surveys 29*(1), 76–96.

Zargar, F. and D. Kumar (2019). Informational inefficiency of bitcoin: A study based on high-frequency data. *Research in International Business and Finance 47*, 344–353.

Appendices

Supplementary Figures and Tables

In Figure A1 we compare end-of-day BTC/USD historical prices from Bitbay, Bitfinex, Bitstamp, Cexio, Exmo, Gemini, ItBit, Kraken, Quadrigacx, Quoinex, and Yobit with the BTC/USD price from the Coinbase exchange, and we plot the empirical density of $\left(p_t^j - p_t^c\right)/p_t^c$, where p_t^j is the end-of-day BTC/USD price from exchange j on day t, and p_t^c is the BTC/USD price in Coinbase at the same time. All end-of-day (23:00:00 UTC) prices are retrieved from Cryptodatadownload for the period 1 January 2018 to 31 December 2018. The average deviation is 1.18% with a standard deviation of 2.58%, but the distribution exhibits a strong positive skew and heavy tails, with some deviations being as large as 4%.

Figure A1: **Percentage deviations of daily BTC exchange prices relative to Coinbase**
Empirical density of the difference between the BTC/USD end-of-day prices on Bitbay, Bitfinex, Bitstamp, Cexio, Exmo, Gemini, ItBit, Kraken, Quadrigacx, Quoinex, and Yobit, and the BTC/USD end-of-day prices on Coinbase. The differences are expressed relative to Coinbase prices, the data is retrieved from Cryptodatadownload, and the sample period is 1 January 2018 – 31 December 2018.

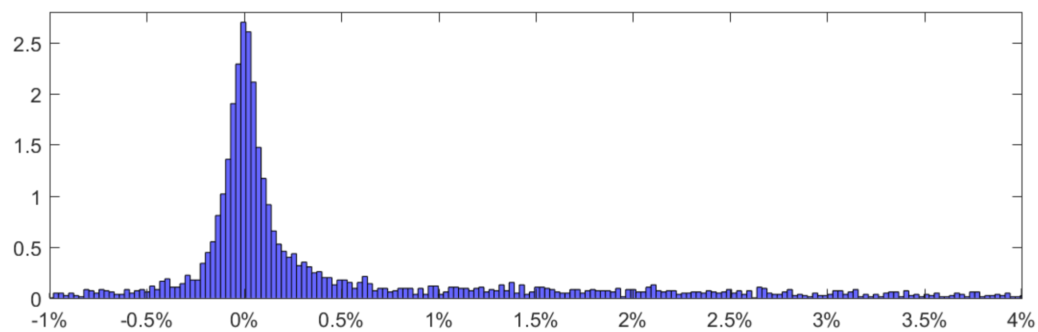


Figure A2 depicts the empirical density of the intra-day BTC price range for the period 1 January 2016 to 31 December 2018 period. The high-low range is calculated using Cryptocompare's BTC/USD index price and is expressed relative to each day's Low price.

Figure A2: **Intra-day range of Cryptocompare BTC/USD**
Empirical density of the intra-day range for Cryptocompare's BTC/USD single index. The range is expressed relative to each day's Low price, and the sample period is 1 January 2016 – 31 December 2018.

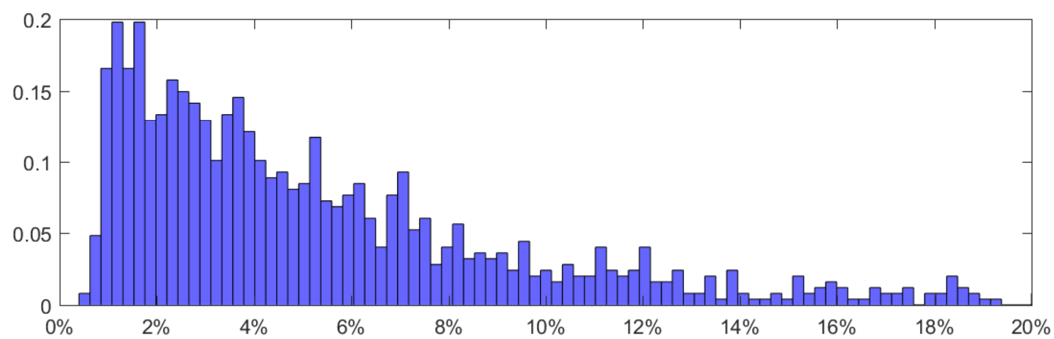


Table A1: **Sample Statistics on BTC Non-Traded Price Indices and Traded Prices**

Sample statistics of the daily returns on BTC prices from CG, CM and CC, and from Bitfinex, Coinbase, Gemini, Kraken and Poloniex.

	BTC Non-Traded			BTC Traded				
	CG	**CM**	**CC**	**Bitfinex**	**Coinbase**	**Gemini**	**Kraken**	**Poloniex**
1 Apr 2016 – 31 Mar 2017								
Mean (p.a.)	113.1%	110.8%	112.6%	111.3%	113.5%	113.1%	113.4%	113.9%
Volatility	58.9%	56.6%	58.0%	55.9%	58.5%	58.3%	60.0%	59.9%
Skewness	-0.74	-0.73	-0.84	-0.58	-0.50	-0.84	-0.79	-0.88
Ex. Kurtosis	6.24	6.01	6.62	5.79	6.03	6.11	6.51	7.09
1 Apr 2017 – 31 Mar 2018								
Mean (p.a.)	241.7%	241.2%	240.4%	243.0%	238.5%	239.7%	238.0%	242.0%
Volatility	105.2%	104.1%	104.3%	107.8%	103.9%	104.7%	102.5%	106.7%
Skewness	0.41	0.33	0.28	0.35	0.46	0.42	0.27	0.34
Ex. Kurtosis	4.57	2.57	2.20	2.06	2.68	2.68	1.97	2.01
1 Apr 2018 – 31 Mar 2019								
Mean (p.a.)	-33.9%	-33.9%	-31.9%	-29.2%	-31.7%	-31.4%	-31.3%	-31.6%
Volatility	61.8%	61.6%	63.8%	65.6%	64.4%	64.6%	64.6%	64.2%
Skewness	-0.15	-0.13	-0.16	-0.07	-0.19	-0.17	-0.18	-0.18
Ex. Kurtosis	2.96	3.02	2.90	3.12	3.20	3.13	3.14	3.20

Table A2: **Sample Statistics on ETH Non-Traded Price Indices and Traded Prices**

Sample statistics of the daily returns on ETH/USD prices from CG, CM and CC, and from Bitfinex, Coinbase, Gemini, Kraken and Poloniex.

	ETH Non-Traded			ETH Traded				
	CG	**CM**	**CC**	**Bitfinex**	**Coinbase**	**Gemini**	**Kraken**	**Poloniex**
1 Jul 2016 – 31 Mar 2017								
Mean (p.a.)	238.8%	247.8%	249.8%	241.7%	252.1%	250.1%	247.9%	249.5%
Volatility	113.4%	114.1%	115.3%	110.0%	118.1%	117.9%	114.6%	115.9%
Skewness	0.97	1.32	1.17	1.49	1.32	0.84	1.15	1.08
Ex. Kurtosis	8.49	8.12	6.87	6.66	6.70	5.38	6.63	6.69
1 Apr 2017 – 31 Mar 2018								
Mean (p.a.)	290.8%	293.8%	292.7%	294.4%	291.6%	298.2%	289.6%	293.2%
Volatility	130.7%	133.6%	132.9%	133.3%	132.7%	136.2%	130.2%	132.4%
Skewness	0.54	0.76	0.66	0.41	0.81	0.79	0.60	0.43
Ex. Kurtosis	2.06	2.74	1.81	1.24	2.37	2.11	1.41	1.16
1 Apr 2018 – 31 Mar 2019								
Mean (p.a.)	-56.9%	-57.5%	-52.2%	-51.3%	-52.7%	-52.5%	-51.9%	-53.9%
Volatility	95.5%	95.0%	99.2%	99.0%	99.5%	99.6%	99.7%	97.9%
Skewness	0.05	-0.04	-0.02	0.02	-0.07	-0.03	-0.04	-0.01
Ex. Kurtosis	2.02	2.08	1.98	1.95	2.05	2.11	2.00	1.99

Table A3: **Optimal GARCH Models – DIC and IC Comparison**

Comparison of DIC (upper metric in each row) and IC (lower metric in each row) for two-state Markov-switching GARCH models. We examine all combinations of GARCH, GJR-GARCH and EGARCH with normal and Student-t (both symmetric and skewed) distributed innovations. We use BTC data from CG, CM, CC, Bitstamp and Kraken, for the sample period 1 September 2013 – 31 March 2019. We highlight (in blue) the optimal (smallest) DIC and IC values for each data source.

	GARCH				**GJR-GARCH**				**EGARCH**			
	Normal		**Student-t**		**Normal**		**Student-t**		**Normal**		**Student-t**	
	Sym.	**Sk.**	**Sym.**	**Sk.**	**Sym.**	**Sk.**	**Sym.**	**Sk.**	**Sym.**	**Sk.**	**Sym.**	**Sk.**
CG	10298.45	10285.59	10253.31	10227.27	10334.05	10330.67	10246.01	10246.49	10317.50	10260.63	10233.19	10233.08
	10304.86	10292.79	10274.62	10238.78	10344.67	10339.49	10259.90	10264.81	10332.53	10271.10	10240.76	10248.11
CM	10274.03	10267.27	10201.66	10188.53	10297.07	10305.19	10253.51	10189.07	10321.82	10320.95	10198.15	10189.12
	10280.70	10273.99	10215.43	10200.89	10303.14	10314.02	10263.73	10198.58	10337.98	10329.20	10205.55	10199.76
CC	10510.80	10498.39	10385.62	10377.72	10514.47	10499.12	10381.61	10382.14	10430.68	10425.66	10371.59	10369.67
	10517.30	10504.90	10393.36	10386.59	10523.79	10507.59	10392.58	10394.86	10437.89	10432.08	10378.40	10381.09
Bitstamp	10362.92	10354.06	10302.68	10286.02	10404.15	10410.56	10316.93	10318.08	10415.23	10410.00	10300.69	10291.61
	10369.28	10360.75	10320.41	10298.88	10412.13	10418.72	10331.79	10330.84	10428.35	10424.19	10308.80	10302.95
Kraken	10518.65	10516.61	10524.27	10486.93	10532.42	10532.82	10462.73	10459.52	10484.96	10483.45	10449.89	10445.04
	10524.56	10523.98	10582.89	10523.98	10539.56	10541.38	10470.79	10467.94	10491.50	10490.66	10458.82	10456.94

Table A4: **Markov Switching GARCH on Non-Traded and Traded BTC Data**

Markov switching skewed Student-t GARCH on BTC daily returns from CG, CM, CC, Bitstamp and Kraken. The sample period is 1 September 2013 – 31 March 2019. The two-state Markov-switching skewed Student-t GARCH model is defined as $\sigma_{it}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{i,t-1}^2$, where $\varepsilon_t \sim t_{\eta_i,\xi_i}$ and $i = 1, 2$. Again we approximate the steady-state volatility as $UV_i = \sqrt{\frac{365\,\omega_i}{1-\alpha_i-\beta_i}}$. We highlight in blue the values of interest.

	CG	**CM**	**CC**	**Bitstamp**	**Kraken**
			State 1		
ω_1	0.164 (0.121, 0.227)	0.066 (0.049, 0.092)	0.174 (0.112, 0.394)	0.065 (0.050, 0.088)	0.002 (0.001, 0.004)
α_1	0.078 (0.065, 0.101)	0.062 (0.050, 0.074)	0.076 (0.066, 0.100)	0.055 (0.045, 0.067)	0.036 (0.027, 0.045)
β_1	0.909 (0.887, 0.923)	0.927 (0.913, 0.939)	0.914 (0.888, 0.926)	0.930 (0.918, 0.940)	0.961 (0.949, 0.971)
η_1	2.450 (2.410, 2.508)	2.412 (2.379, 2.453)	2.374 (2.336, 2.486)	2.455 (2.417, 2.496)	2.339 (2.297, 2.396)
ξ_1	1.042 (1.015, 1.065)	1.051 (1.031, 1.073)	1.027 (0.985, 1.051)	1.050 (1.028, 1.072)	1.117 (1.092, 1.141)
p_{11}	0.966 (0.959, 0.972)	0.965 (0.958, 0.971)	0.976 (0.973, 0.979)	0.958 (0.951, 0.965)	0.935 (0.920, 0.951)
UV_1	68.2%	48.6%	82.1%	40.3%	14.8%
			State 2		
ω_2	3.550 (2.213, 5.457)	1.674 (1.236, 2.336)	5.925 (4.401, 7.112)	1.633 (1.229, 2.152)	2.086 (1.357, 3.470)
α_2	0.241 (0.186, 0.296)	0.161 (0.135, 0.193)	0.330 (0.249, 0.385)	0.158 (0.138, 0.186)	0.182 (0.147, 0.238)
β_2	0.703 (0.621, 0.777)	0.823 (0.786, 0.854)	0.588 (0.539, 0.657)	0.829 (0.797, 0.853)	0.805 (0.745, 0.846)
η_2	4.275 (3.988, 4.536)	3.857 (3.689, 4.048)	3.516 (3.280, 3.699)	4.143 (3.959, 4.330)	3.543 (3.319, 3.759)
ξ_2	0.856 (0.826, 0.893)	0.870 (0.845, 0.894)	0.911 (0.880, 0.955)	0.869 (0.844, 0.895)	0.975 (0.948, 0.999)
p_{22}	0.959 (0.952, 0.966)	0.964 (0.959, 0.969)	0.976 (0.972, 0.979)	0.960 (0.954, 0.965)	0.943 (0.931, 0.955)
UV_2	153%	198%	163%	215%	242%

Table A5: **Markov Switching EGARCH on Non-Traded and Traded BTC Data**

Markov switching skewed Student-t EGARCH on BTC daily returns from CG, CM, CC, Bitstamp and Kraken. The sample period is 1 September 2013 – 31 March 2019. The two-state Markov-switching skewed Student-t EGARCH model is defined as $\ln(\sigma^2_{it}) = \omega_i + g(z_{t-1}) + \beta_i \ln(\sigma^2_{i,t-1})$, where $g(z_t) = \theta_i z_t + \gamma_i(|z_t| - \mathrm{E}[|Z_t|])$, $Z_t \sim t_{\eta_i,\xi_i}$ and $i = 1, 2$. Again we approximate the steady-state volatility as $UV_i = \sqrt{365 \exp(\frac{\omega_i}{1-\beta_i})}$. We highlight in blue the values of interest.

	CG	**CM**	**CC**	**Bitstamp**	**Kraken**
			State 1		
ω_1	-0.029 (-0.034, -0.025)	0.001 (-0.002, 0.005)	0.003 (-0.001, 0.006)	-0.009 (-0.013, -0.006)	0.000 (-0.001, 0.001)
γ_1	0.207 (0.186, 0.232)	0.376 (0.312, 0.446)	0.284 (0.239, 0.334)	0.195 (0.176, 0.215)	0.119 (0.084, 0.171)
θ_1	0.050 (0.036, 0.065)	0.040 (-0.008, 0.085)	0.110 (0.069, 0.154)	0.074 (0.060, 0.089)	0.101 (0.080, 0.120)
β_1	0.995 (0.995, 0.996)	0.995 (0.995, 0.995)	0.994 (0.993, 0.994)	0.991 (0.991, 0.992)	0.996 (0.996, 0.997)
η_1	2.944 (2.849, 3.050)	2.100 (2.100, 2.100)	2.166 (2.160, 2.173)	2.869 (2.793, 2.978)	2.305 (2.283, 2.329)
ξ_1	1.030 (1.012, 1.050)	1.044 (1.027, 1.062)	1.039 (1.019, 1.062)	1.031 (1.009, 1.051)	1.091 (1.072, 1.111)
p_{11}	0.878 (0.867, 0.888)	0.976 (0.974, 0.978)	0.971 (0.969, 0.973)	0.895 (0.887, 0.905)	0.965 (0.963, 0.968)
UV_1	0.871%	21.3%	23.5%	11.1%	19.9%
			State 2		
ω_2	0.569 (0.472, 0.677)	0.294 (0.246, 0.365)	0.420 (0.360, 0.492)	0.464 (0.366, 0.547)	1.207 (1.010, 1.502)
γ_2	0.373 (0.306, 0.445)	0.353 (0.300, 0.408)	0.473 (0.418, 0.531)	0.321 (0.255, 0.387)	0.381 (0.301, 0.459)
θ_2	-0.045 (-0.083, -0.006)	0.015 (-0.010, 0.040)	-0.063 (-0.096, -0.033)	-0.055 (-0.086, -0.019)	-0.132 (-0.173, -0.091)
β_2	0.866 (0.830, 0.891)	0.919 (0.897, 0.933)	0.884 (0.865, 0.902)	0.895 (0.869, 0.918)	0.652 (0.570, 0.710)
η_2	9.669 (8.757, 10.665)	4.434 (4.156, 4.711)	3.827 (3.690, 3.988)	14.412 (12.849, 15.971)	5.720 (5.291, 6.170)
ξ_2	0.768 (0.733, 0.804)	0.837 (0.810, 0.867)	0.883 (0.855, 0.910)	0.691 (0.660, 0.727)	0.994 (0.964, 1.024)
p_{22}	0.633 (0.622, 0.645)	0.967 (0.966, 0.968)	0.965 (0.964, 0.966)	0.637 (0.626, 0.649)	0.942 (0.939, 0.944)
UV_2	158%	116%	118%	173%	108%