Where is the information on USD/Bitcoin hourly price movements?

Article *in* Notas Económicas · December 2017

DOI: 10.14195/2183-203X_45_1

3 authors, including:

Helder Sebastiao
University of Coimbra

António Portugal Duarte
University of Coimbra

Where is the information on USD/Bitcoins hourly price movements?

HELDER SEBASTIÃO

CeBER and Faculty of Economics of the University of Coimbra

ANTÓNIO PORTUGAL DUARTE

CeBER and Faculty of Economics of the University of Coimbra

GABRIEL GUERREIRO

CeBER and Faculty of Economics of the University of Coimbra

CeBER Working Papers

No. 5 2017

Where is the information on USD/Bitcoins hourly price movements?

Helder Sebastião[*], António Portugal Duarte, and Gabriel Guerreiro

CeBER, Faculty of Economics, University of Coimbra

April, 2017

This paper analyses the price discovery process in the USD/Bitcoin market since the Mt.Gox bankruptcy until the aftermath of the hack attack on Bitfinex (01-Mar-2014 until 30-Nov-2016). The Geweke feedback measures, estimated pairwise using hourly returns, show that there is a positive relationship between the total feedback and market share, measured by trading volume, that most of the information is transmitted between exchanges within an hour, at least for the main four exchanges (Bitfinex, Bitstamp, BTC-e and ItBit), while lagged feedback runs mainly from the major exchange. Other minor exchanges seem to react to price information with some delay and are thus considered as merely satellite exchanges. Bitfinex stands out as the most important exchange in transmitting information to the market: the relative importance of the lagged feedback from Bitfinex to the market is 18.29% while the lagged feedback from the market to Bitfinex accounts only for 0.60% of the total feedback. The volatility in the major exchange in each pair is the main factor explaining the feedback measures, sustaining the claim that the information-based component of volatility increases with the relative dimension of the exchange.

JEL classification: F13, G12, G14, G15.
Keywords: Bitcoin, price discovery, Geweke feedback measures, volume, volatility.

[*]Corresponding author. Postal address: Faculdade de Economia, Universidade de Coimbra, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal. Phone number: +351 239 790 570, fax number: +351 239 790 514, e-mail address: helderse@fe.uc.pt.

1. Introduction

Bitcoin is a decentralised open source peer-to-peer (P2P) crypto-currency protocol. The Bitcoin project was firstly presented in a self-published paper by Nakamoto (2008)[1] on October 31, 2008, and it was announced in the cryptography mailing list on January 11, 2009. Nakamoto's paper describes a mathematical system that could be used to produce and manage a virtual currency, mainly designed for supporting online transactions. Its main merit, which is the basis for its success in relation to other virtual currencies, is to solve the double spending problem (when an individual, conducting an online transaction, sends the same money to two counterparts at the same time) without the need for a third trusted intermediary. Moreover, while other online payment systems, such as Paypal, eBay, Dwolla, WePay, Google Wallet and WePay, still have impediments in cross-border transactions, Bitcoin allows its holders to trade across borders, in an increasingly global marketplace (ECB, 2012; Lancelot and Tatar, 2013; Pagliery, 2014; Pieters, 2016).

As a crypto-currency, Bitcoin is entirely digital, without physical existence nor country of origin. Bitcoin is issued and controlled by its users and is accepted among the members of an increasing virtual community, therefore is not subjected to any regulation or supervision from a monetary authority. Bitcoins are created (discovered) by solving a complex mathematical algorithm in a process known as "mining", which is transparent, decentralised, and overseen by the Bitcoin protocol users. The Bitcoin supply has increased at a predictable rate, depending on the number of "miners" and traders, technological advances and energy costs, however Bitcoin tends to be subjected to a deflationary process as the demand becomes higher than the supply (Nakamoto, 2008; Fink and Johann, 2014).

Bitcoins are sent and received via Bitcoin addresses. However, because there is no central processing authority, transactions between users must be confirmed by consensus: a private Bitcoin key of one user has to match the public Bitcoin key of another user. This is made possible through the Bitcoin's "blockchain", which is essentially a public chronological log of every confirmed Bitcoin transaction (ECB, 2012).

The historical appreciation of Bitcoin has been absolutely impressive. This can be grasped by some anecdotal evidence: the first product bought using Bitcoins was two pizzas on May

[1] The identity of Satoshi Nakamoto is currently unknown. The name is a pseudonym of someone who worked initially on the Bitcoin project but only interacted with people in developer forums. At the end of 2010, Satoshi Nakamoto disappeared from these forums, announcing his departure and handing off the project to the open source community. No one knows his (hers or their) true identity, but it is said that this entity retains approximately 100 million USD worth of Bitcoins at the 2013 prices. For more details, see Bradbury (2013) and Velde (2013).

21, 2010, for a price of 10000 Bitcoins, roughly 25 USD at that time (Fink and Johann, 2014). At the time of writing, the price for one Bitcoin is around 1188.46 USD; so, at the actual prices, this is probably the most expensive meal in the history of mankind! The exponential appreciation of Bitcoin seems to be behind the increasing interest that Bitcoin is gaining in the online trading community.

Since its online creation in 2009, Bitcoin has grown from a new digital currency traded essentially between enthusiasts, to a booming monetary system receiving substantial media attention for its conceptual merits. The market capitalisation of Bitcoin surpassed 19 billion USD recently, and the transaction volume keeps growing in a more global and diversified scale. Approximately 16.26 million Bitcoins are currently (April 7, 2017) in circulation and there are more than fifty Bitcoin exchanges offering trades against different currencies, such as AUD, CAD, CHF, CNY, EUR, GBP, HKD, JPY, PLN and USD.

The days when a single exchange completely dominated the market are gone and currently several exchanges compete for trading volume. During its first year, Bitcoins were traded solely privately, however, in 2010, the first currency exchanges emerged, with Mt.Gox claiming the market leadership. Throughout the next two years, Mt.Gox kept its leader position, holding a market share of more than 80% (Brandvold et al., 2015). Later, exchange trading volumes at Bitstamp, BTC-e and Bitfinex rose as Mt.Gox's fell down, due to several technical incidents and legal issues. In the latter half of 2013, those three exchanges took more than 50% of USD/BTC market share. Mt.Gox suspended all transactions in February 2014 after a serious security breach, not without before Bitcoin price increased by a factor of 10, roughly from 100 USD to 1000 USD.

In terms of economic literature, the study of the Bitcoins phenomenon is still relatively limited, namely in respect to the price discovery process on the currency exchanges. This paper addresses this issue by examining transaction data on fourteen Bitcoin exchanges that were active at least one year during the sample period, namely Bitfinex, Bitstamp, BTC-e, Coinbase, ItBit, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBX, BitKonan, and Bitbay.

Our main aim is to analyse the Bitcoin exchange market since the Mt.Gox bankruptcy (March 1, 2014) until the aftermath of the hack attack on Bitfinex (November 30, 2016). We intend to study low trading frequency exchanges, and, for this reason, we choose a sampling interval of one hour. In order to assess the informational relationship between these fourteen Bitcoin exchanges, we apply the feedback measures proposed by Geweke (1982). We also conduct a second stage analysis using panel regressions of the feedback measures on market variables, such as volatility and volume.

Besides using data on a significant number of Bitcoin exchanges, which includes not only the biggest USD/BTC exchanges but also a diverse set of minor exchanges, this study also contributes to the literature by carrying out an investigation on a sample beginning with the Mt.Gox bankruptcy, while, to the best of our knowledge, most of the studies on Bitcoin price dynamics focused on the period until that event. Hence, this paper is the first one to describe short run price dynamics in the USD/BTC exchange market for such a recent sampling period.

This study shows that there is a positive relationship between the total feedback and market share, measured by trading volume, and that most of the information is transmitted between exchanges within an hour, at least for the main four exchanges (Bitfinex, Bitstamp, BTC-e and ItBit). Lagged feedback runs mainly from the major exchange to the other exchanges, being its relative importance positively related to the difference in trading volume. Bitfinex stands out as the most important exchange in transmitting information to the USD/BTC exchange market. The minor exchanges seem to react to price information with some delay, thus confirming that they act as merely "satellite exchanges".

The remainder of the paper is organised as follows. Section 2 describes some of the main characteristics and mechanisms associated with the Bitcoin transaction system. Section 3 presents a brief literature review. Section 4 refers to the data and a preliminary analysis. Section 5 describes the methodology, namely the Geweke feedback measures and the procedure for the panel regression analysis. The results of the empirical application are presented in Section 6. The paper concludes in Section 7.

2. Characteristics and mechanisms of the Bitcoin transaction system

As mentioned earlier, Bitcoin is a decentralised open source peer-to-peer (P2P) electronic cash system that can be used to purchase goods and services on the internet in the same way that printed currencies are used as a medium of exchange. However, contrary to "real cash", Bitcoin is a virtual currency without any formal organisational structure behind and is not backed up by any central monetary authority or country. Instead, Bitcoin transaction system is managed entirely by a P2P network that manages balances and electronic transactions on its own.

Those features are possible through the historical record of all transactions that occur between Bitcoin addresses. An individual dealing in Bitcoins can use any number of addresses. For the transaction to happen, a private Bitcoin key of one user has to match the public Bitcoin key of another user. When a Bitcoin is sent from one address to another, the transaction is collected by the network and because there is no central processing authority, transactions must be confirmed by mutual agreement. Trades are collected into logical entities known as blocks. All transactions are anonymously added to the Bitcoin database, the blockchain citepECB, that is, blocks of new transactions are being added chronologically to the chain. Unlike other digital currencies, with Bitcoin the accounts' addresses are pseudonymous and the protocol is designed to encourage the use a new address for each transaction. Using cryptography Bitcoin creates mathematical proofs to secure the blockchain and safeguard users' transactions. These proofs are verified by miners who process the blocks as part of the blockchain (Bradbury, 2013; Lancelot and Tatar, 2013; Fink and Johann, 2014).

The mining process validates all previous transactions and generates new Bitcoins through the use of widely distributed software that endlessly attempts to solve the complex mathematical problem securing the blockchain. The mining process is consequently a transparent and decentralised activity, allowing the system to transmit transactions to the network almost immediately and verify them within an hour.

The P2P protocol ensures that a single ordering of transactions becomes apparent and becomes accepted by all. For example, if a user transfers Bitcoins, the operation is firstly added to a pool of pending transactions, then miners verify it by solving the mathematical puzzle and the transaction is added to the blockchain, thus becoming part of the list of all historical transactions. Additionally, all users verify the correctness of the blockchain to ensure that double spending or any kind of fraud is prevented (Barber et al., 2012; Fink and Johann, 2014). In this way, users contribute directly to the integrity of the Bitcoin system. Mining activities require considerable CPU power, which results in significant expenditures in energy and time (Lancelot and Tatar, 2013). Therefore, for every block processed miners are rewarded with the creation of new Bitcoins, which ensures that the Bitcoins supply increases at a predictable growth rate. Nowadays, a new block is "won" approximately every 10 minutes, meaning that it takes on average about 10 minutes for a miner or mining pool to solve the cryptographic problem. The winning miner is awarded a given amount of Bitcoins (currently 25 BTC) while the losers get nothing, accordingly this mining activity is characterized as a "competitive bookkeeping" (Harvey, 2016).

Bitcoin was projected to be used as a medium of exchange worldwide, supplied at a predictable growth rate, until a deterministic maximum of 21 million Bitcoins. The larger the Bitcoin community and the total computational resources devoted to Bitcoin generation, the more difficult the computational problem becomes. Under these circumstances, the Bitcoin is subjected to a deflation phenomenon as it seems that the Bitcoin usage increases relatively more than the generation of new Bitcoins (ECB, 2012). Hence, it is expected that the real purchasing power of Bitcoins will increase over time and users, most probably, will hold on to their Bitcoins rather than spent them.

Bitcoin's invention is nevertheless revolutionary, since for the first time in monetary history and online trading activity the well-known double spending problem, a situation in which an individual conducting an online transaction attempts to send the same money to two users at once, is solved without the need for a third trusted part (bank or other financial intermediary). This innovative feature of Bitcoin results in economic advantages, even in relation to the existing "real" currencies, such as low transfer costs, absence of intermediary costs and anonymity (Barber et al., 2012).

But there are also some disadvantages. Velde (2013) and Fink and Johann (2014), for instance, argue that Bitcoins are stored in digital wallets, which are prone to hacker attacks, and that there is no deposit insurance as in the case of fiat money issued by banks. There is no central clearing nor third part guaranteeing activity. Under these circumstances, no monetary policy can be conducted, for example, to promote economic growth. Additionally, short selling Bitcoins is almost impossible to carry out, which in practice is a severe drawback to Bitcoin to act as an economical financing instrument. The brief history of Bitcoin is replete of events of malpractices, money laundry and financing criminal activities, hence no wonder it is regarded with suspicion by regulators and monetary authorities. For instance, the ECB (2012) does not exclude the possibility of being in the presence of a speculative bubble or even a Ponzi scheme, in view of the still relatively limited Bitcoin's use as currency, as well as to its high volatility and operational risk.

Nevertheless, Bitcoin is increasingly gaining a relevant economic status, coexisting with the current fiat money systems. In fact, most Bitcoin users do not only engage in mining activities, but also trade Bitcoins against "real" currencies. The online currency exchanges are important components of the cryptocurrency ecosystem as they connect Bitcoin with real economy, in which transactions are denominated in local currencies (Li and Wang, 2017).

3. Literature review

Economic literature on the Bitcoins issue is quite limited so far. Most papers and books on Bitcoin are from the fields of computer sciences and cryptography, therefore focusing essentially on the explanation of technical and methodological features of the Bitcoin network, mining activity and blockchain knowledge.[2]

Barber et al. (2012), Eyal and Sirer (2014) and Böhme et al. (2015) discuss technical aspects of the Bitcoin project. The authors try to understand what made the Bitcoin transaction system so successful and how Bitcoin could become a good candidate for a long-lived virtual currency, while decades of research on cryptographic e-cash has not lead to a large-scale deployment. Tu and Meredith (2015) and Karame et al. (2015) analyse security and legal issues in crypto-currency systems, while Reid and Harrigan (2013) and Ron and Shamir (2013) dedicate more attention to the analytical aspects related to the information contained in the blockchain. The latter authors show in particular that a large fraction of issued Bitcoins is "dormant", in the sense that they were issued and never traded again. In fact, just about 13% of all Bitcoin transactions are exchange traded (Fink and Johann, 2014).

An issue that has also attracted some attention in the academic world is the discussion on if Bitcoin can be considered in fact a currency, particularly in view of its still relatively low acceptance in the foreign exchange market and its poor performance as a medium of store of value. Naturally, central banks have been quite concerned with this issue, namely the ECB (2012) argues that, like any currency, Bitcoin depends on trust. However, compared to a commodity-backed currency, as in the international gold standard system or fiat money, the trust is not supported by its intrinsic value or on the belief in a central monetary authority solvency, but rather on cryptography and computer technology. Although several concepts of money have been associated to the Bitcoin phenomenon, such as "crypto-currency" (Elias, 2011; Evans, 2014; Böhme et al., 2015), "digital currency" (Grinberg, 2011; Dwyer, 2015) or "virtual currency" (ECB, 2012; Tu and Meredith, 2015), for some authors Bitcoin cannot be considered a currency. Yermack (2013), for instance, argues that the Bitcoin exhibits excess volatility, has no correlation with classical currencies and is not regulated. Brière et al. (2015) also argue that Bitcoins seem to be a valuable asset for portfolio diversification. Following a similar line of thought, Fink and Johann (2014) defend that, in its current usage, Bitcoin is more an investment vehicle than a currency. This conclusion arises from the observation that investors' motives in trading Bitcoins are purely speculative and that more and more uniformed investors are trying to participate in this type of investment.

A few set of studies have also investigated the Bitcoin exchange market. Some of these studies focus their attention on the existence of speculative bubbles. Cheung et al. (2015), for instance, observe several short-lived bubbles and three huge bubbles in the period between 2011 and 2013. Cheah and Fry (2015) also show that Bitcoin exhibits speculative bubbles and

[2] Velde (2013) presents a comprehensive overview of the Bitcoin project. A literature review on Bitcoin can be found in Li and Wang (2017) whose general structure we follow in this section.

argue that the fundamental price of Bitcoin is surprisingly equal to zero. In this context, it is not surprising that authors, such as Glaser et al. (2014), have questioned the motivations behind the implementation of Bitcoin and the resemblance of its exchange activities to pure speculative trading.

More recently, the economic literature on Bitcoin was directed predominantly towards the conduction of econometric analyses regarding the identification and explanation of the main determinants of the Bitcoin exchange rate.

Kristoufek (2013) shows a very high correlation between the Google Trends measure, the number of Wikipedia views on Bitcoins and the Bitcoin exchange rate. The author suggests that trading in Bitcoins is attention-driven and uniformed and that it resembles the speculation activity during the early 2000 internet bubble. In a later study, Kristoufek (2015) uses a wavelet coherency analysis to classify the correlation between Bitcoin exchange rate and various determinants, such as speculative behaviour and the exchange-trade ratio. The results show that these determinants play a significant role at lower frequencies. Using an Autoregressive Distributed Lag (ARDL) approach, Bouoiyour and Selmi (2015) also identify several determinants of the Bitcoin exchange rate, including Google searches, hash rate, ratio of exchange-trade volume and stock market dynamics. Conducting a similar analysis, Polasik et al. (2015) also analysed the impact on the Bitcoin monthly returns of several variables: the number of new Bitcoin articles, Google searches, transaction volume, number of Bitcoins and a set of economic determinants, such as industrial production growth, inflation and unemployment. The authors conclude that Bitcoin returns are mainly driven by news volume, news sentiment and the number of traded Bitcoins.

In a quite different perspective, Pieters (2016) assumes that the Bitcoin market is mainly driven by information and proposes a method that uses Bitcoin daily prices to construct unofficial exchange rates of several currencies, detect the size of capital controls to the official exchange rates and identify exchange rate regimes. As noticed by Pieters (2016) , because Bitcoin prices are directly observed, this method requires no reporting agents, thereby removing potential reporting bias. Additionally, because Bitcoin may be used alternatively as an investment vehicle, varying Bitcoin exchange rates exist for both floating and managed currencies. Pieters (2016) also document that Bitcoin-based exchange rates reveal that there is no consistent pattern in Granger causality between unofficial and official rates in different regimes or capital controls and that countries probably engage in short-interval capital controls.

In what concerns to price discovery process in Bitcoin currency exchanges, to the best of our knowledge, the existing literature is quite scarce. Nevertheless, the econometric techniques used and the results achieved in these studies allow us to gather an important set of results, fundamental for contextualizing our own study.

Fink and Johann (2014) are one of the first studies on the price formation and market microstructure of the Bitcoin. Volatility, turnover, liquidity, returns, price efficiency, and price cointegration, are all investigated in detail. Furthermore, the Bitcoin ownership structure and its implications on those characteristics is also analysed. The authors shows that Bitcoin prices experience extreme returns and high volatility. With some surprise, the authors find that the Bitcoin price is not informationally efficient. Market fragmentation and liquidity increased in recent years, and the largest Bitcoin exchanges are cointegrated. Transaction frequency, ownership, and size are broadly dispersed across more than fifteen million Bitcoin users. This dispersion shows that the Bitcoin is traded by both retail and professional traders with different strategies. In fact, although the Bitcoin network is large, it is dominated by a small number of big exchanges that own and trade a high fraction of the available Bitcoins. The price discovery leader was the Mt.Gox exchange before its bankruptcy, but after that event the market shares and price discovery across Bitcoin exchanges are more balanced. Brandvold et al. (2015) also study the contributions of several Bitcoin exchanges to the price discovery process. Since it is expected that a higher fraction of the price discovery process of Bitcoin occurs at exchanges with higher trading volumes, the authors analyse the ratio between the information share and the activity share. The authors conclude that for the whole sample period (April 1, 2013 – February 25, 2014) Mt.Gox, together with the BTC-e, were the market leaders, while the rest of the exchanges were less informative, but still providing some information to the Bitcoin exchange market. They also determine that information shares are dynamic and evolving significantly over time. While Mt.Gox dominated the price discovery process, its information share decreased significantly but still was higher than its activity share. BTC-e was one of the most informative exchanges and was much more informative than other exchanges during the big shock to the Bitcoin market resulting from the shut-down of the Silk Road.

4. Data and preliminary analysis

The data for this study was mainly collected from the site www.bitcoincharts.com. This aggregation site compiles transaction data on several exchanges that trade bitcoins against different currencies, being the USD and CNY the most important ones. Although Bitcoin high frequency data is available for free in other public sites, it seems that this database is quite reliable and has already been used in several academic papers (see, for instance, Fink and Johann, 2014;

Brandvold et al., 2015; Pieters, 2016).

In this paper we just focus on the USD/BTC market. The main reason for this relies on the fact that there has been some rumours that the main exchanges dealing with the Chinese Yuan, which of course, have their headquarters in China, tend to exaggerate their trading volume in order to attract more traders.[3] Even for the USD/BTC market this issue creates a data reliability problem: according to data available at data.bitcoinity.org, OkCoin, the second biggest Chinese Bitcoin exchange following BTC China has, during the sample period, a 6.91% share in the USD/BTC market, measured by traded volume in BTC.[4] However, it seems that transaction data on BTC against USD at OkCoin is not publicly available.

The sample period was defined by two particular events. On February 25, 2014, Mt.Gox closed permanently for business. Before its bankruptcy, Mt.Gox was by far the dominant exchange in the USD/BTC market with a share of 74.83% of trading volume (January 2010 until February 2014). Even at the closing day, the daily market share of Mt.Gox (33.46%) was still above the market share of any of its rivals (Bitstamp 28.01%, Bitfinex 20.60%, and BTC-e 17.12%). On the early afternoon of August 2, 2016, Bitfinex halted trading after discovering that roughly 120 thousand BTCs were stolen, and, on that day, in just a few hours, the Bitcoin dropped nearly 15% (from 560.16 USD at 19:30 UTC to 480 USD at 24:00 UTC), as the overall market reacted to news on the hack.[5] Bitfinex stayed closed for seven days, until August 8, 2016. On December 8, 2017, the site bitcoincharts ended publishing Bitfinex data. Given these events and the data available we selected a sample period of 1006 days, since 01/03/2014 until 30/11/2016.

Given the availability of transaction data, we also had to decide on the sampling frequency. There is a trade-off between gathering as much information as one can and avoiding the effects of microstructural noise and non-synchronous trading. For instance, Fink and Johann (2014) use a 1-minute interval while Brandvold et al. (2015) use a 5 minutes interval. Here, because we intend to study also low trading frequency exchanges we choose a sampling interval of one hour. At this frequency we collect information on hourly price indexes, weighted by trading volume, and trading volume in BTC. The use of price indexes instead of transaction prices (e.g. last price before the sampling point) smooths the price time series and diminishes the impact

[3]For instance, the total traded volume of BTC against CNY, since March 15, 2015, until March 14, 2017, according to the site data.bitcoinity.org was approximately 1.3 billion, which roughly means a market share of 94% during that period, while the USD market share was only 4.06%. About this issue see, for instance, the news article "Chinese Bitcoin Exchange OKCoin Accused of Faking Trading Data", written by Eric Mu on December 21, 2013 (available at: http://www.coindesk.com).

[4]https://data.bitcoinity.org/.

[5]http://www.coindesk.com.

of extreme trades documented in Brandvold et al. (2015). On the other hand, it allows us to take into account that Bitcoin may be traded at really small quantities. One Bitcoin can be divided down to one satoshi, i.e. 10^{-8} of a unit, and, in fact, trades with volumes lower than 0.1 BTC are the most common ones (Brandvold et al., 2015) .

Finally, we had to decide what exchanges to use in this study that trade USD/BTC and have data available at the bitcoincharts site. The criterion was to consider those exchanges that were active at least one year during the sample period (01/03/2014 to 30/11/2016). We end up with 14 exchanges, which account for 74.34% of the total Bitcoins traded against USD during the sample period (72.71 million BTC). Table 1 presents some information on these exchanges with a focus on its trading activity.

Table 1: Exchange information

This table shows some information on the 14 exchanges used in this study, namely: Headquarters, period of data availability at www.bitcoincharts.com, total trading volume USD/BTC in millions of BTC (where the values in parenthesis present the volume of each exchange relative to the total trading volume of the overall USD/BTC market - 72.71 million BTC according to https://data.bitcoinity.org), trade duration in minutes and seconds and average volume per trade during the sampling period (01-Mar-2014 to 30-Nov-2016).

Exchanges	**Headquarters**	**Data Available**	**Volume**	**Trade Duration**	**Volume per trade**
Bitfinex	Hong Kong	Full sample	22.148 (30.46%)	10s	2.548
Bitstamp	Luxembourg	Full sample	11.099 (15.27%)	12s	1.532
BTC-e	Bulgaria	Full sample	7.3712 (10.14%)	5s	0.424
Coinbase	San Francisco USA	Since 01/12/2014	5.0439 (6.94%)	7s	0.406
ItBit	New York USA	Full sample	3.6011 (4.95%)	1m59s	4.930
LakeBTC	Shanghai China	Until 19/06/2015	2.1103 (2.90%)	24s	0.583
LocalBitcoins	Finland	Since 11/03/2013	1.6223 (2.23%)	52s	0.971
Kraken	San Francisco USA	Full sample	0.4260 (0.59%)	3m11s	0.936
HitBTC	UK	Full sample	0.3526 (0.48%)	1m36s	0.389
Onecoin	Bulgaria	09/03/2014-04/04/2015	0.2318 (0.31%)	1m54s	0.029
Rock	Malta	Full sample	0.0206 (0.03%)	23m34s	0.335
CampBX	Atlanta USA	Until 19/10/2016	0.0150 (0.02%)	36m49s	0.267
BitKonan	Croatia	Full sample	0.0096 (0.01%)	58m6s	0.385
Bitbay	Poland	Since 16/05/2014	0.0091 (0.01%)	19m12s	0.121

Bitfinex, Bitstamp and BTC-e stand out as the three main exchanges with a total volume of roughly 75% of the USD/BTC market formed by these 14 exchanges, and in a second level are Coinbase and ItBit with roughly 15% of the total volume. In order to analyse the price discovery process among all exchanges we need a fairly continuous time series without many gaps, hence we decide to isolate Bitfinex, Bitstamp, BTC-e and ItBit from all the other exchanges. Therefore we create a pool of exchanges, which we denominate by "Others", by combining the price and volume information on Coinbase, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBx, BitKonan and Bitbay. The exchange Coinbase is included into this basket not because its trading volume is low but due to its late opening on the 01-Dec-2014, nine months after the sample beginning. The trading volume for Others is simply obtained by adding up the trading volume of these 10 exchanges, while the price is computed as an average of prices in these exchanges, using the trading volume as a weighting scheme.

From now on we will assume that the USD/BTC market was totally composed, since 01-Mar-2014 until 30-Nov-2016, by Bitfinex with a market share, given by the relative trading volume of 40.97%, Bitstamp, 20.53%, BTC-e, 13.64%, ItBit, 6.66%, and Others, 18.20%. In this last case, it means an average market share per exchange of 1.82%. Table 2 shows the preliminary statistics of the hourly logarithmic returns for the exchanges under scrutiny.

The number of staled prices seems only to be a problem for ItBit, where 14.8% of the returns series is zero. The mean and median returns are almost zero, but the returns show positive and negative extreme values. This is particularly true for ItBit, with a minimum and a maximum hourly returns of -50.56% and 54.28%, and for the basket Others with a minimum of -47.71% and a maximum of 49.67%. The standard deviation is inversely related with the exchange's dimension; for instance, the standard deviation of Others is more than twice the standard deviation of the four bigger exchanges. The returns are obviously non-normal, presenting negative skewness (except for the ItBit) and leptokurtosis. ItBit also shows a higher kurtosis than the other exchanges. The first order autocorrelations are significantly positive, except for Others that is negative. The second and third order autocorrelations are all significantly negative (except the third order autocorrelation for Others). Although we expected that persistence would increase with lower trading intensity and that it should be higher in Others as result of the averaging procedure, the Bayesian-Information Criterion indicates that modelling the returns of Others by an autoregressive process would imply using a 52 lag length, which means using self-information for more than two days.

Finally we verified if all return series were stationary by applying Augmented Dickey-Fuller (ADF) tests, without constant and trend, and with a lag length inferred by the Bayesian-Information Criterion (BIC). For all the returns series the tests were categorical in rejecting the null hypotheses of a unit root at a 1% significance level.[6]

[6]The statistics are -80.0 for Bitfinex, -79.7 for Bitstamp, -79.3 for BTC-e, -70.2 for ItBit and -21.9 for Others.

Table 2: Descriptive statistics on returns

This table summarises the statistics for the hourly logarithmic returns of the USD/BTC exchange rates. The sample covers the period since 01/03/2014 until 30/11/2016, for a total of 1006 days (24143 hourly observations). The exchanges are Bitfinex, Bitstamp, BTC-e, ItBit and "Others". This last one refers to a compilation of several minor exchanges (Coinbase, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBx, BitKonan and Bitbay). The Others' price upon which are computed the returns is the price index averaged by volume. BIC denotes the Bayesian-Information Criterion for choosing the lag length in an autoregressive process. The autocorrelations significance levels were inferred using Bartlett's standard errors. Values significant at the 10%, 5% and 1% levels are marked by *, ** and ***, respectively.

	Bitfinex	**Bitstamp**	**BTC-e**	**ItBit**	**Others**
No. of zeros	203	108	653	3580	0
	(0.8%)	(0.4%)	(2.7%)	(14.8%)	
Mean (10^{-5})	1.2531	1.3131	1.3203	1.1261	1.0927
Minimum	-0.1656	-0.1390	-0.1498	-0.5056	-0.4771
Percentile 5	-0.0086	-0.0087	-0.0080	-0.0082	-0.0257
Median	0.0000	0.0001	0.0000	0.0000	0.0003
Percentile 95	0.0085	0.0083	0.0080	0.0079	0.0248
Maximum	0.1053	0.1178	0.1016	0.5428	0.4967
Stand. deviation	0.0063	0.0062	0.0062	0.0081	0.0179
Skewness	-1.0749	-0.6081	-0.9125	2.1412	-0.0693
Kurtosis	43.590	34.074	48.419	151.76	68.504
Jarque-Bera (10^6)	1.6620***	0.9728***	2.0786***	2307.6***	4.3163***
Autocorr(1)	0.1282***	0.1416***	0.1139***	0.0089	-0.3578***
Autocorr(2)	-0.0879***	-0.0772***	-0.0589***	-0.2176***	-0.0187***
Autocorr(3)	-0.0466***	-0.0488***	-0.0365***	-0.0154**	0.0092
BIC	3	3	3	5	52

5. Methodology

In order to assess the informational relationship between exchanges we use the feedback measures of Geweke (1982). These measures are applied pairwise for each pair of exchanges and between each exchange and the rest of the market, formed by all the other relevant exchanges. We also proceed with a second stage analysis by conducting panel regressions of the feedback measures on market variables, such as volatility and volume.

Consider that a pair of time series of returns, $\{r_{it}, r_{jt}\}$, sampled at a given frequency, lets say hourly, can be expressed as a bivariate autoregressive process of an arbitrary order p:

$$\begin{bmatrix} r_{it} \\ r_{jt} \end{bmatrix} = \begin{bmatrix} A(L) & B(L) \\ C(L) & D(L) \end{bmatrix} \begin{bmatrix} r_{it} \\ r_{jt} \end{bmatrix} + \begin{bmatrix} \epsilon_{it} \\ \epsilon_{jt} \end{bmatrix}, \tag{1}$$

where $A(L)$, $B(L)$, $C(L)$ and $D(L)$ are polynomials in the lag operator, L, and the innovations are Gaussian (i.e. ϵ_{kt} are independently and identically $N(0, \sigma_k^2)$, for $k = i, j$). The innovations covariance matrix is

$$\mathbf{\Omega} = cov \begin{bmatrix} \epsilon_{it} \\ \epsilon_{jt} \end{bmatrix} = \begin{bmatrix} \sigma_i^2 & \sigma_{ij} \\ \sigma_{ji} & \sigma_j^2 \end{bmatrix}. \tag{2}$$

It is also assumed that the innovations are serially uncorrelated, but they may be correlated with each other contemporaneously and at various leads and lags.

Absence of Granger causality, denoted by "$\tilde{\rightarrow}$", implies that the coefficient matrix is triangular in the VAR representation. For a bivariate process there are two no lagged feedback null hypotheses: $H_{i\tilde{\rightarrow}j} : C(L) = 0$ and $H_{j\tilde{\rightarrow}i} : B(L) = 0$. Under these hypotheses, the VAR simplifies to:

$$\begin{bmatrix} r_{it} \\ r_{jt} \end{bmatrix} = \begin{bmatrix} A(L) \\ D(L) \end{bmatrix} \begin{bmatrix} r_{it} \\ r_{jt} \end{bmatrix} + \begin{bmatrix} \xi_{it} \\ \xi_{jt} \end{bmatrix}. \tag{3}$$

Additionally, if there is no contemporaneous linear relationship between the series, $H_{i\tilde{\leftrightarrow}j}$, then $cov(\xi_{it}, \xi_{jt}) = 0$. The hypothesis of no linear link between the two variables is given by the conjunction of the previous hypotheses: $H_{i\tilde{\cdot}j} = H_{i\tilde{\rightarrow}j} \cap H_{i\tilde{\leftrightarrow}j} \cap H_{j\tilde{\rightarrow}i}$. The measures proposed by Geweke (1982) allow to test these hypotheses:

Measure of lagged feedback from i to j:

$$F_{i \to j} = ln \left(\frac{\sigma^2_{\xi_j}}{\sigma^2_{\epsilon_j}} \right). \tag{4}$$

Measure of lagged feedback from j to i:

$$F_{j \to i} = ln \left(\frac{\sigma^2_{\xi_i}}{\sigma^2_{\epsilon_i}} \right). \tag{5}$$

Measure of contemporaneous feedback between i and j:

$$F_{i \leftrightarrow j} = ln \left(\frac{\sigma^2_{\epsilon_i} \sigma^2_{\epsilon_j}}{|\mathbf{\Omega}|} \right). \tag{6}$$

Measure of total feedback between i and j:

$$F_{i.j} = ln \left(\frac{\sigma^2_{\xi_i} \sigma^2_{\xi_j}}{|\mathbf{\Omega}|} \right). \tag{7}$$

Where $|\mathbf{\Omega}|$ denotes de determinant of the innovations covariance matrix in the unrestricted model. Under the null hypothesis these measures, multiplied by the number of observations, T, are asymptotically independent and follow chi-squared distributions with p, p, 1 and $2p + 1$ degrees of freedom, respectively.

The feedback measures are just the log-likelihood ratio statistics for the null hypotheses, and, therefore, if feedback is present, their asymptotic distributions are well defined. Under the alternative hypothesis, these measures, multiplied by the number of observations, are asymptotically non-central chi-squared:

$$T\hat{F}_{i \to j} \sim \chi^{'2}(p, TF_{i \to j}), \tag{8}$$

$$T\hat{F}_{j \to i} \sim \chi^{'2}(p, TF_{j \to i}), \tag{9}$$

$$T\hat{F}_{i \leftrightarrow j} \sim \chi^{'2}(1, TF_{i \leftrightarrow j}), \text{and} \tag{10}$$

$$T\hat{F}_{i.j} \sim \chi^{'2}(2p + 1, TF_{i.j}). \tag{11}$$

The Geweke feedback measures have several advantages over other methodologies, such as the Wald F-test: (i) under the alternative hypothesis these statistics represent cardinal measures of the extent of linear dependence in the two series; (ii) these measures are additive: $F_{i.j} = F_{i\to j} + F_{i\leftrightarrow j} + F_{j\to i}$; (iii) the comparison between the feedback in two pair of variables is straightforward as long as the measures are estimated using the same number of observations; (iv) these metrics are unaffected by prefiltering the series by any invertible lag operator (Parzen, 1982), which suggests that they are less sensitive to the effects of non-synchronous trading and other microstructural idiosyncratic sources of noise; and (v) feedback measures are well-defined and readily interpreted in the presence of additional regressors, i.e. they can be interpreted as conditional linear feedback measures (Geweke, 1984).

In the second stage of our analysis, we compute the time series of the feedback measures for each different pair of exchanges using a non-overlapping rolling window with constant length. This rolling window procedure is also used to compute the time series of the trading intensity, measured by the log-volume in bitcoins, *vol*, and of the volatility for each exchange. Although volume and volatility are usually highly correlated, they may account for different types of information arrival processes (Andersen, 1996). For measuring volatility we use the range estimator of Parkinson (1980):

$$HL = \left[\left(\frac{1}{D}\right) \sum_{d=1}^{D} \frac{\left(\ln\left(\frac{H_d}{L_d}\right)\right)^2}{4\ln(2)} \right]^{\frac{1}{2}}, \tag{12}$$

where D is the number of days in the window and H_d and L_d are the maximum and minimum prices recorded on day d. Although the Parkinson estimator assumes no drift and it tends to underestimate volatility, it seems a good candidate to measure volatility in a continuous trading market (other more efficient range volatility estimators, such as Garman and Klass, 1980; Rogers and Satchell, 1991; Yang and Zhang, 2000, also consider the opening and closing prices).

The regression analysis was conducted as follows. Firstly, the feedback measures were normalized using the procedure prescribed by Geweke (1982): If $T\hat{F} \sim \chi'^2(df, TF)$ where df is the degree of freedom and TF is the non-centrality parameter, then

$$n\hat{F} = (|T\hat{F} - (df-1)/3|)^{1/2} \sim N((|TF - (df-1)/3|)^{1/2}, 1). \tag{13}$$

Secondly, for each pair of exchanges, i and j, and for each normalized measure, $nF_{i\to j}$, $nF_{j\to i}$, $nF_{i\leftrightarrow j}$ and $nF_{i.j}$, we construct a matrix of regressors, $[HL_i, HL_j, vol_i, vol_j,]$. With the purpose of simplifying the interpretation of the results, the pair (i, j) is built considering in the first entry the exchange with the highest market share (measured by trading volume). So, for N exchanges we have $N(N-1)/2$ time series on each feedback measure. Finally, for each feedback measure we run the following panel regression:

$$n\hat{F}_{(i,j)t} = \beta_0 + \beta_1 HL_{it} + \beta_2 HL_{jt} + \beta_3 vol_{it} + \beta_4 vol_{jt} + \vartheta_{(i,j)t}. \qquad (14)$$

The regression analysis on the feedback measures has been used elsewhere. For instance, Kawaller et al. (1993) use this methodology to study the interrelationship between stock index and stock index futures, Bracker et al. (1999) study the evolution of integration, measured by contemporaneous feedback, between several national stock markets. In this last paper, the authors use a pooled regression and combine the two lagged feedback measures, arguing that they are analogous in economic and statistical terms.

Our perspective is different. Firstly, we do not superimpose the data pooling and instead let the data tell us what is the best model (pooled regression, panel with fixed effects or panel with random effects). However, because we have a fixed and relatively small set of units of interest (the pairs of exchanges) there is a presumption in favour of fixed effects. Secondly, we do not aggregate the lagged feedback measures and instead we model them separately. Obviously, the two measures are statistically similar but they may be economically different, the impact of volatility and volume from a particular exchange on a feedback measure may be different depending on the exchange position in the pair (i, j), i.e., if it is the leader or the follower exchange, in terms of market share.

The analysis design allows us to formulate several hypotheses. Basically, most of these hypotheses are drawn upon the Wall Street adage "It takes volume to make prices move". On other hand, we also assume that volatility is mostly information-driven, especially if it is from the leader exchange and therefore volatility should increase the exchanges' proximity.

From the pairwise estimations of the feedback measures we can test the following hypothesis:

> $\mathbf{H_1}$: *The ranking of the pairs of exchanges by the total feedback is the same as its ranking by the combined volume of the two exchanges.*
>
> $\mathbf{H_2}$: *At an hourly sampling, the great contributor for the total feedback is the contemporaneous feedback, and its contribution increases with the combined volume of the two exchanges.*
>
> $\mathbf{H_3}$: *In each pair, the lagged feedback runs mostly from the exchange with higher volume to the other exchange, and the difference between the lagged feedbacks is positively related to the difference in trading volumes.*

In the same line of reasoning, we can also formulate hypotheses on the expect signs of the regressors in Eq. (14).

> $\mathbf{H_4}$: *All the variables in the contemporaneous feedback regression have positive signs.*
>
> $\mathbf{H_5}$: *All the variables in the total feedback regression have positive signs.*
>
> $\mathbf{H_6}$: *In the lagged feedback regressions, $i \to j$, volume and volatility of exchange i have positive signs, while volume and volatility of exchange j have negative signs.*

In the next section we present the empirical results that allow us to infer about the validity of these hypotheses.

6. Results

Firstly, we estimate the feedback measures pairwise, considering the exchanges Bitfinex, Bitstamp, BTC-e, ItBit and Others, where this last one is a pool of minor exchanges (Coinbase, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBx, BitKonan and Bitbay). The estimates were obtained from fitting VARs with a lag structure truncated at lag 52, which is the longest lag structure inferred by the Bayesian-Information Criterion applied to the univariate time series of hourly continuous returns. Using such lag length enable us to capture all the autocorrelation and lagged cross-correlation structure, even in the Others returns. Results are presented in Table 3.

As expected, the total feedback is highly correlated with the average market share, implying that the interrelationship between exchanges increases with their relative volume. However, the ordering isn't exactly the same and the total feedbacks between Bitstamp and BTC-e and between Bitstamp and ItBit are higher than the total feedback between Bitfinex and ItBit, despite this last pair sharing a higher trading volume. This probably means that market proximity, in terms of trading volume, also tightens prices together.

The contemporaneous feedback is the main contributor to the total feedback, except when Others is included in the pair. In this case, the contemporaneous feedback only accounts for about to 34% to 39% of the total feedback, and most of the feedback runs from the major exchange to Others (52.78% to 64.70%). The contemporaneous feedback ranges from 96.11% of the total feedback in the Bitfinex/Bitstamp pair and only 34.61% of the total feedback in the Bitfinex/Others pair. The lagged feedback is asymmetrical and runs dominantly from the

Table 3: Pairwise estimation of feedback measures

Geweke´s feedback measures were estimated for all pairs of exchanges using hourly logarithmic returns. The column "Average Share" gives the total market share of the exchanges divided by the number of exchanges (2 for all pairs, except for the pairs that include Others, where the divisor is 11). The "Average Share" is used to order the pairs in the table. Others refers to a compilation of several minor exchanges (Coinbase, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBx, BitKonan and Bitbay). The feedback measures were obtained from fitted VAR models with a lag structure truncated at 52. The lagged feedback from i to j and from j to i are denoted by $F_{i\to j}$ and $F_{j\to i}$, respectively, while the simultaneous feedback is denoted by $F_{i\leftrightarrow j}$ and the total feedback is $F_{i.j}$. The relative weight (i.e. divided by the total feedback) of the lagged feedbacks and simultaneous feedback are shown in parentheses. All the estimates are significant at the 1% level, except the lagged feedback from Others to Bitfinex and to Bitstamp that are not significant at the 10% level. These two estimates are marked by ([a]).

Exch.(i)	***Exch.(j)***	**Average Share**	$F_{i\to j}$	$F_{i\leftrightarrow j}$	$F_{j\to i}$	$F_{i.j}$
Bitfinex	**Bitstamp**	30.75%	0.0495	1.3948	0.0070	1.4512
			(3.41%)	(96.11%)	(0.47%)	
Bitfinex	**BTC-e**	27.31%	0,059	0.8783	0.0089	0.9463
			(6.24%)	(92.82%)	(0.94%)	
Bitfinex	**ItBit**	23.82%	0.1361	0.4485	0.0039	0.5886
			(23.13%)	(76.21%)	(0.67%)	
Bitstamp	**BTC-e**	17.09%	0.0412	0.8838	0.0220	0.9469
			(4.35%)	(93.33%)	(2.32%)	
Bitstamp	**ItBit**	13.60%	0.1211	0.4830	0.0071	0.6112
			(19.81%)	(79.02%)	(1.17%)	
BTC-e	**ItBit**	10.15%	0.0888	0.3447	0.0128	0.4463
			(19.90%)	(77.24%)	(2.87%)	
Bitfinex	**Others**	5.38%	0.1615	0.0864	0.0017 ([a])	0.2497
			(64.70%)	(34.61%)	(0.69%)	
Bitstamp	**Others**	3.52%	0.1653	0.0974	0.0018 ([a])	0.2645
			(62.51%)	(36.82%)	(0.67%)	
BTC-e	**Others**	2.89%	0.1300	0.0811	0.0038	0.2149
			(60.49%)	(37.74%)	(1.77%)	
ItBit	**Others**	2.26%	0.1033	0.0758	0.0167	0.1958
			(52.78%)	(38.70%)	(8.52%)	

major exchange than the other way around, these figures range from 0.0495 (3.41% of total feedback) in the Bitfinex/Bitstamp pair to 0.1615 (64.70%) in the Bitfinex/Others pair. The feedback from the minor exchange is quite marginal, with a maximum absolute value of 0.022 in the Bitstamp/BTC-e pair and a maximum relative value of 8.52% in the ItBit/Others pair. In fact, the only estimates that are not significant (even at the 10% level) are the lagged feedback from Others to the two major exchanges, Bitfinex and Bitstamp.

Overall, Table 3 indicate that the three major markets are highly integrated. In these markets, the relative contemporaneous feedbacks estimates suggest that more than 92% of price variability is communicated between markets within one hour. The level of integration decays with ItBit, which has a relative contemporaneous feedback of around 77% with the three major markets. The basket Others mostly reacts to price changes with a delay of at least one hour and therefore the minor exchanges compiled into Others may be seen as pure satellite exchanges, in the sense of Garbade and Silber (1979). However we have to keep in mind that this last result is in part due to smoothing the price series across ten minor exchanges.

Although the results suggest that Bitfinex is the dominant market in terms of the transmission of short run price information, we now try to answer directly to this question. In order to position each exchange in the overall USD/BTC market we computed the feedback measures between each exchange and the Market, where its return is computed upon the price index averaged by volume of the remaining exchanges. Table 4 presents these results, where in Panel A the Market includes the basket Others and Panel B considers the Market formed only by the most important four exchanges.

Not surprisingly, we notice that when we exclude Others from the analysis, all the lagged feedback measures from an exchange to the Market decrease, while all the lagged feedback measures from the Market to an exchange increase. The degree of integration (contemporaneous feedback) is quite higher when minor exchanges are excluded, which also roughly doubles the total feedback. One can observe from Panel A that the contemporaneous feedback is the major contributor to the total feedback, with this measure presenting a relative weight above 61%, except in the case Others/Market, where this figure only reaches 34.73%.

The feedback from the Market to Others is quite high (63.32%) while the inverse lagged feedback is marginal (1.95%). Moreover, when we include Others in the Market, the lagged feedback from the Market to Bitfinex is not significant at the 10% level. This corroborates the previous conclusion that Others doesn't has, on average, important information on the USD/BTC price movements. Given these results, we hereafter study the USD/BTC market formed only by Bitfinex, Bitstamp, BTC-e and ItBit.

Table 4, Panel B, deserves special attention. The four exchanges are well integrated, with

Table 4: Feedback measures between each exchange and the market

Geweke´s feedback measures were estimated for all pairs exchange/Market, using hourly logarithmic returns. For the Market, denoted by M, the returns are computed upon the price index, weighted by volume, of all remaining exchanges. Panel A includes in the Market the minor exchanges compiled into the basket Others, while Panel B only considers Bitfinex, Bitstamp, BTC-e and ItBit. The feedback measures are estimated from fitted VAR models with a lag structure truncated at 52. The lagged feedback from i to M and from M to i are denoted by $F_{i\rightarrow M}$ and $F_{M\rightarrow i}$, respectively,while the simultaneous feedback is denoted by $F_{i\leftrightarrow M}$ and the total feedback is $F_{i.M}$. The relative weight (i.e. divided by the total feedback) of the lagged feedbacks and simultaneous feedback are presented in parentheses. All the estimates are significant at the 1% level, except the lagged feedback from the Market to Bitfinex that is not significant at the 10% level. This estimate is marked by ([a]).

Exchange (i)	$F_{i\rightarrow j}$	$F_{i\leftrightarrow j}$	$F_{j\rightarrow i}$	$F_{i.j}$
		Panel A: Including Others		
Bitfinex	0.2031	0.3219	0.0020 ([a])	0.5270
	(38.55%)	(61.08%)	(0.40%)	
Bitstamp	0.1666	0.4779	0.0082	0.6527
	(25.53%)	(73.21%)	(1.26%)	
BTC-e	0.1122	0.3786	0.0151	0.5060
	(22.18%)	(74.83%)	(2.99%)	
ItBit	0.0565	0.2953	0.0737	0.4256
	(13.27%)	(69.40%)	(17.33%)	
Others	0.0048	0.0849	0.1549	0.24454
	(1.95%)	(34.73%)	(63.32%)	
		Panel B: Excluding Others		
Bitfinex	0.1702	0.7542	0.0060	0.9304
	(18.29%)	(81.06%)	(0.65%)	
Bitstamp	0.1020	0.8824	0.0255	1.0099
	(10.10%)	(87.37%)	(2.53%)	
BTC-e	0.0401	0.7701	0.0407	0.8510
	(4.72%)	(90.50%)	(4.79%)	
ItBit	0.0086	0.4465	0.1300	0.5852
	(1.47%)	(76.31%)	(22.21%)	

more than 75% of the information on prices being transmitted to the overall market within an hour. The feedback from Bitfinex and from Bitstamp to the Market is higher than the reverse feedback, while the opposite happens for BTC-e and ItBit. However, it takes more than an hour for transmitting 18.29% of the short run price movements that have its origin in Bitfinex, while the relative lagged feedback from Bitstamp is 10.10%. The feedback from the Market to Bitfinex and to Bitstamp is only 0.60% and 2.55%, respectively. In sum, one might say that Bitstamp is more integrated with the overall market, but Bitfinex has the short run informational dominance.

The above outcomes are drawn upon the overall sample period and therefore may be interpreted as unconditional results. Now we analyse how the feedback measures relate to volatility and volume. As described in the methodological section we partitioned the sample into non-overlapping rolling windows and estimate the time series of feedback measures. We choose a window with an amplitude of 5 days, which means that the VARs estimates are obtained from subsamples of 119 returns observations.[7] For the sample from 01/03/2014 until 30/11/2016, 1006 days, we get 201 estimates for each feedback measures.

Table A1 in appendix presents some statistics on the time series of the feedback measures. It is interesting to notice the asymmetry between the significance of the estimates of the lagged feedbacks. The lagged feedback from the major exchange to the minor exchange, $F_{i\to j}$, is significant at least 50% of the times and reaches a maximum of 82.41% in the Bitfinex/ItBit pair, while the lagged feedback from the minor exchange to the major exchange, $F_{j\to i}$, is not significant most of the times, and in the case of the Bitfinex/BTC-e pair it is only significant in 15.58% of the times. The cross-correlation between the feedback measures highlights the proximity of the simultaneous and total feedback ($corr\left(F_{i\leftrightarrow j}, F_{i.j}\right) = 0.9776$) and that the lagged feedback from the major exchange is negatively correlated with the simultaneous feedback ($corr\left(F_{i\to j}, F_{i\leftrightarrow j}\right) = -0.0963$) and naturally with the other lagged feedback from the minor exchange ($corr\left(F_{i\to j}, F_{j\to i}\right) = -0.1426$).

The summary statistics of volatility and log-volume are shown in Table A2, also in appendix. The volatility series present similar means and standard deviations, around 0.02, positive skewness and above normal kurtosis, especially for the ItBit volatility. Hence, the Jarque-Bera test rejected the normality of all the series. The log-volume series present a mean ranging from 9.6564 for the Bitfinex to 7.7818 for the ItBit. Standard deviations are quite lower than the means, ranging from 0.5836 for the BTC-e to 1.0505 for the ItBit. The Jarque-Bera test does not reject the normality of log-volume for the Bitfinex and Bitstamp, however the negative

[7]We also conduct the regression analysis using rolling windows of 15 days. The results although consistent with the ones presented here were less robust, arguably due to averaging the variables across more days.

skewness and high kurtosis of log-volume of BTC-e and ItBit imply the non-normality of these series. The contemporaneous cross-correlation between volatility and log-volume is quite high, around 0.38, as expected, but still it puts into perspective the existence of intrinsic features of each series.

Before we proceed with the panel regressions we test for unit roots in the series, with special attention to the log-volume series, and conclude for its stationarity.[8] Then we select the panel model using the Breusch-Pagan test on the null hypothesis of pooled regression and the Hausman test on the null hypothesis of consistency of the GLS estimates (random effects). All the statistics were significant at the 1% level, which led us to select the panel regression with fixed effects.[9] The estimation results are shown in Table 5.

The joint significance test indicates the model adequacy for all the feedback measures at a 1% significance level, however the within coefficient of determination is quite low for the lagged feedbacks (around 2%), while the R^2 for the contemporaneous feedback is 49.21% and the R^2 for the total feedback is 51.69%. The regressors in the lagged feedback from the major exchanges, $nF_{i\to j}$, have all the expected signs, however only the volatility in exchanges i is significant at the 1% level. The variables of exchanges j are significant at the 5% level. In the equation of the lagged feedback from the minor exchanges, $nF_{j\to i}$, the volatility in exchanges i has the expected sign and is significant at the 5% level, while volume in exchanges j has the expected sign and is significant at the 10% level. These results suggest that the main driving force behind the lagged feedback is the volatility in the major markets, extending the information transmission for more than an hour from the major exchanges and diminishing the reverse feedback. The regression results for the contemporaneous and total feedbacks are quite similar. In fact, the main difference is that the coefficients and the t-statistics in the total feedback are slightly lower. In these two regressions all the variables are significant at the 1% level, and both volumes and volatility in the major exchanges contribute positively for the contemporaneous and total feedback. However, volatility in the minor exchanges has a negative sign implying that an increase in that volatility tends to diminish market integration (contemporaneous feedback) and the total linear interconnection between exchanges.

[8] We use the Im-Pesaran-Shin panel unit root test, with a constant, without trend or lags. The t-bar statistics were the following: $nF_{i\to j}$ -12.445, $nF_{i\leftrightarrow j}$ -7.3391, $nF_{j\to i}$ -12.785, $nF_{i.j}$ -7.5031, HL_i -8.4515, HL_j -9.5308, vol_i -5.9687, vol_j -7.5344, which are all significant at the 1% level.

[9] The Breusch-Pagan statistics for the normalized feedback measures, $nF_{i\to j}$, $nF_{i\leftrightarrow j}$, $nF_{j\to i}$ and $nF_{i.j}$ were 75.219, 3558.1, 188.74 and 5564.8, respectively; while the Hausman statistics were 49.721, 289.60, 81.979 and 387.99, respectively.

Table 5: Panel regressions on the feedback measures

This table shows the panel regression results on the normalized feedback measures, namely the parameter estimates, the Arellano (2003) t-statistics adjusted for heteroscedasticity and serial correlation (in parentheses), the F test for joint significance of the "named regressors" and the within R^2. The feedback measures were estimated, from fitted VAR models with a lag structure truncated at 5, for all pairs of exchanges, using hourly logarithmic returns for each sub-sample of 5 days. The feedback measures, multiplied by the number of observations, were then normalised. The panel regressions consider 201 time points for 6 cross-section units (each pair of exchanges) for a total of 1206 observations. The normalized lagged feedback from i to j and from j to i are denoted by $nF_{i\to j}$ and $nF_{j\to i}$, respectively, while the simultaneous feedback is denoted by $nF_{i\leftrightarrow j}$ and the total feedback is $nF_{i.j}$. Values significant at the 10%, 5% and 1% levels are marked by *, ** and ***, respectively.

Variables	$nF_{i\to j}$	$nF_{i\leftrightarrow j}$	$nF_{j\to i}$	$nF_{i.j}$
Intercept	4.090***	-21.091***	1.0296	-16.415***
	(6.3539)	(-8.6080)	(1.7039)	(-7.5586)
HL_i	18.442***	81.589**	-6.4511**	76.866**
	(11.909)	(3.9798)	(-3.9287)	(3.8838)
HL_j	-16.228**	-65.268***	4.7885	-57.726***
	(-3.9841)	(-5.2993)	(1.6733)	(-4.5083)
vol_i	0.0720	1.5370***	0.0070	1.3934***
	(0.8073)	(4.6319)	(0.3557)	(4.6017)
vol_j	-0.1735**	1.9216***	0.1583*	1.6458***
	(-3.2619)	(6.0128)	(2.4148)	(5.6185)
$F_{(4,5)}$	75.058***	654.77***	16.883***	2026.9***
R^2	0.0234	0.4921	0.0215	0.5169

7. Conclusion

The present paper aims to analyse the price discovery process among all relevant exchanges in the USD/Bitcoin market with public available data, even those with low trading intensity. The data was collected from the site www.bitcoincharts.com and reflects the trading information on 14 exchanges for the period since the Mt.Gox bankruptcy until the aftermath of the hack attack on Bitfinex, i.e., since 01-Mar-2014 until 30-Nov-2016, for a total of 1006 days (24143 hourly observations). Given the traded volume and the period of trading, we decided to study Bitfinex, Bitstamp, BTC-e and ItBit separately, while aggregating the remaining 10 exchanges (Coinbase, LakeBTC, LocalBitcoins, Kraken, HitBTC, Onecoin, Rock, CampBx, BitKonan and Bitbay) into a basket, that we denominated by Others. The aggregating procedure uses the price index weighted by trading volume.

The Geweke feedback measures were then estimated pairwise between exchanges, using hourly returns (computed on price indexes weighted by volume) for the all sampling period. The results highlight the existence of a positive relationship between the total feedback and market share of both exchanges but also with its proximity in terms of trading volume. Most of the information is transmitted between exchanges within an hour, at least for the main four exchanges, while lagged feedback runs mainly from the major exchanges in each pair, being its relative importance positively related to the difference in trading volumes. The minor exchanges, compiled into the pool Others, seem to react to price information with some delay and are merely satellite exchanges.

The Geweke feedback measures were also estimated pairwise between each exchange and the rest of the market. The results supported the main conclusions stated above, namely that the consideration of minor exchanges only brings more noise into the price index process, Bitstamp is well integrate with the overall market, but, more importantly, Bitfinex stands out as the most important exchange in transmitting information to the market: the relative importance of the lagged feedback from Bitfinex to the market is 18.29% while that quantity for the lagged feedback from the market to Bitfinex is only 0.60%.

The panel regression of the feedback measures on volatility and volume show that these variables explain a fair part of the contemporaneous and total feedback, with all the signs being significantly positive except the volatility in the minor exchange. This result suggests that pairwise, in relative terms, the volatility in the major exchange is mainly information-based, aligning exchanges together, while volatility in the minor exchange is more noise-based, driving exchanges apart. For the lagged feedback, the most important explaining variable is the volatility in the major exchange which has an obvious different impact: an increase in that volatility increases the feedback from the major exchange while decreases the feedback from the minor exchange.

Trading Bitcoins involves an important operational risk (the history of Bitcoin exchanges is replete of events such as hack attacks, missing wallets, malpractices, government interventions, temporary and not so temporary trading halts, etc.) and the market industrial organization is in permanent evolution. Therefore, our results are quite conditional on the sampling period. In fact, it would be quite interesting to see if the informational superiority of Bitfinex still exists after the hack happened on August 2016.

References

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *The Journal of Finance*, 51(1):169–204.

Arellano, M. (2003). *Panel data econometrics*. Oxford University Press.

Barber, S., Boyen, X., Shi, E., and Uzun, E. (2012). Bitter to better - How to make Bitcoin a better currency. In *Financial Cryptography and Data Security*, pages 399–414. Springer. 16th International Conference, FC 2012, Keromytis A. D. (eds.).

Böhme, R., Christin, N., Edelman, B., and Moore, T. (2015). Bitcoin: Economics, technology, and governance. *The Journal of Economic Perspectives*, 29(2):213–238.

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

Bracker, K., Docking, D. S., and Koch, P. D. (1999). Economic determinants of evolution in international stock market integration. *Journal of Empirical Finance*, 6(1):1–27.

Bradbury, D. (2013). The problem with Bitcoin. *Computer Fraud & Security*, 2013(11):5–8.

Brandvold, M., Molnár, P., Vagstad, K., and Valstad, O. C. A. (2015). Price discovery on Bitcoin exchanges. *Journal of International Financial Markets, Institutions and Money*, 36:18–35.

Brière, M., Oosterlinck, K., and Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with Bitcoins. *Journal of Asset Management*, 16(6):365—-373.

Cheah, E.-T. and Fry, J. (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, 130:32–36.

Cheung, A., Roca, E., and Su, J.-J. (2015). Crypto-currency bubbles: An application of the Phillips–Shi–Yu (2013) methodology on Mt.Gox Bitcoin prices. *Applied Economics*, 47(23):2348–2358.

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, 17:81–91.

ECB (2012). *Virtual currency schemes.* European Central Bank, Frankfurt. Available at https://www.ecb.europa.eu/pub/pdf/other/.

Elias, M. (2011). Bitcoin: Tempering the digital ring of gyges or implausible pecuniary privacy. Available at SSRN: https://ssrn.com/abstract=1937769.

Evans, D. S. (2014). Economic aspects of Bitcoin and other decentralized public-ledger currency platforms. Coase–Sandor Institute for Law & Economics Working Paper No. 685.

Eyal, I. and Sirer, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. In *International Conference on Financial Cryptography and Data Security*, pages 436–454. Springer.

Fink, C. and Johann, T. (2014). Bitcoin markets. Available at SSRN: https://ssrn.com/abstract=2408396.

Garbade, K. D. and Silber, W. L. (1979). Dominant and satellite markets: A study of dually-traded securities. *The Review of Economics and Statistics*, 61(3):455–460.

Garman, M. B. and Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 53(1):67–78.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313.

Geweke, J. F. (1984). Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association*, 79(388):907–915.

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C., and Siering, M. (2014). Bitcoin - asset or currency? Revealing users' hidden intentions. In *Twenty Second European Conference on Information Systems*. Tel Aviv.

Grinberg, R. (2011). Bitcoin: An innovative alternative digital currency. *Hastings Science & Technology Law Journal*, 4(1):160–206.

Harvey, C. R. (2016). Cryptofinance. Available at SSRN: https://ssrn.com/abstract=2438299.

Karame, G. O., Androulaki, E., Roeschlin, M., Gervais, A., and Čapkun, S. (2015). Misbehavior in Bitcoin: A study of double-spending and accountability. *ACM Transactions on Information and System Security (TISSEC)*, 18(1):2:1–2:32.

Kawaller, I. G., Koch, P. D., and Koch, T. W. (1993). Intraday market behavior and the extent of feedback between S&P 500 futures prices and the S&P 500 index. *Journal of Financial Research*, 16(2):107–121.

Kristoufek, L. (2013). Bitcoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, 3(3415).

Kristoufek, L. (2015). What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. *PloS ONE*, 10(4):e0123923.

Lancelot, R. and Tatar, J. (2013). *What's the deal with Bitcoins?* People Tested Publications.

Li, X. and Wang, C. A. (2017). The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. *Decision Support Systems*, 95:49–60.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Available at: https://bitcoin.org/bitcoin.pdf.

Pagliery, J. (2014). *Bitcoin: And the future of money*. Triumph Books.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65.

Parzen, E. (1982). Measurement of linear dependence and feedback between multiple time series: Comment. *Journal of the American Statistical Association*, 77(378):320–322.

Pieters, G. C. (2016). Does Bitcoin reveal new information about exchange rates and financial integration? Globalization and Monetary Policy Institute Working Paper 202. Available at SSRN: https://ssrn.com/abstract=2896642.

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

Reid, F. and Harrigan, M. (2013). An analysis of anonymity in the Bitcoin system. In *Security and Privacy in Social Networks*, pages 197–223. Springer.

Rogers, L. C. G. and Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 1(4):504–512.

Ron, D. and Shamir, A. (2013). Quantitative analysis of the full Bitcoin transaction graph. In *International Conference on Financial Cryptography and Data Security*, pages 6–24. Springer.

Tu, K. V. and Meredith, M. W. (2015). Rethinking virtual currency regulation in the Bitcoin age. *Washington Law Review*, 90(271):271–347.

Velde, F. (2013). Bitcoin: A primer. *Chicago Fed Letter*, (Dec).

Yang, D. and Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *Journal of Business*, 73(3):477–492.

Yermack, D. (2013). Is Bitcoin a real currency? An economic appraisal. Technical Report 19747, National Bureau of Economic Research. Available at: http://www.nber.org/papers/w19747.

Appendix

Table A1: Summary statistics on the feedback measures

This Table summarises some statistics for the Geweke feedback measures computed with a non-overlapping rolling window of 5 days. The exchanges under scrutiny are Bitfinex, Bitstamp BTC-e and ItBit. The sample covers the period since 01/03/2014 until 30/11/2016, for a total of 1006 days, i.e., 201 sub-samples of 5-days (except the first estimates that are obtained from a window of 6 days). The feedback measures were estimated from VARs with a maximum lag of 5 hours (119 observations for each estimation). For each pair the last row presents the number of significant estimates at the 1% level and its relative number (i.e., divided by 201) in parentheses. The critical values for the lagged, contemporaneous and total feedbacks (multiplied by the number of observations) are $\chi^2_{0.99}(5) = 15.086$, $\chi^2_{0.99}(1) = 6.6349$ and $\chi^2_{0.99}(11) = 24.725$, respectively.

Exchange(i)/Exchange(j)	$F_{i\rightarrow j}$	$F_{i\leftrightarrow j}$	$F_{j\rightarrow i}$	$F_{i.j}$
Bitfinex/Bitstamp				
Mean	0.1186	1.2783	0.0562	1.4531
Minimum	0.0109	0.0028	0.0013	0.0564
Median	0.1078	1.2365	0.0504	1.4358
Maximum	0.4062	2.8959	0.1761	2.9599
Stand. deviation	0.0725	0.6089	0.0344	0.6038
Significant (%)	145 (72.86%)	198 (99.50%)	52 (26.13%)	199 (100%)
Bitfinex/BTC-e				
Mean	0.1200	0.7651	0.0496	0.9347
Minimum	0.0105	0.0000	0.0065	0.0306
Median	0.1086	0.6990	0.0415	0.8890
Maximum	0.4679	2.2788	0.1829	2.4051
Stand. deviation	0.0734	0.4882	0.0324	0.4877
Significant (%)	151 (75.81%)	195 (97.99%)	31 (15.58%)	197 (98.00%)
Bitfinex/ItBit				
Mean	0.1638	1.0286	0.0559	1.2483
Minimum	0.0000	0.0000	0.0000	0.0000
Median	0.1393	1.0365	0.0465	1.2418
Maximum	0.6343	2.9220	0.2388	3.0094
Stand. deviation	0.1083	0.6995	0.0392	0.6801
Significant (%)	164 (82.41%)	193 (96.98%)	44 (22.11%)	195 (97.99%)
Bitstamp/BTC-e				
Mean	0.0921	0.7233	0.0653	0.8807
Minimum	0.0034	0.0004	0.0046	0.0371
Median	0.0821	0.6944	0.0564	0.8792
Maximum	0.2833	2.1429	0.2821	2.2061
Stand. deviation	0.0554	0.4640	0.0422	0.4641
Significant (%)	111 (55.78%)	193 (96.98%)	64 (32.16%)	195 (97.99%)
Bitstamp/ItBit				
Mean	0.1277	1.0113	0.0794	1.2185
Minimum	0.0000	0.0000	0.0000	0.0000
Median	0.1045	0.9550	0.0700	1.1189
Maximum	0.5340	3.1985	0.4419	3.2658
Stand. deviation	0.0971	0.6824	0.0605	0.6609
Significant (%)	132 (66.33%)	192 (96.48%)	88 (44.22%)	196 (98.49%)
BTC-e/ItBit				
Mean	0.1035	0.5522	0.0826	0.7384
Minimum	0.0000	0.0000	0.0000	0.0000
Median	0.0764	0.4495	0.0705	0.6507
Maximum	0.5598	1.9695	0.4073	2.1206
Stand. deviation	0.0889	0.4175	0.0622	0.4187
Significant (%)	101 (50.75%)	186 (93.47%)	89 (44.72%)	194 (97.49%)

Table A2: Descriptive statistics on volatility and volume

This table summarises the statistics for the 5-days volatility obtained by the Parkinson range estimator and the log-volume in BTC. The sample covers the period since 01/03/2014 until 30/11/2016, for a total of 1006 days, i.e., 201 observations. BIC denotes the Bayesian-Information Criterion for choosing the lag length in an autoregressive process. The autocorrelations significance level was inferred using Bartlett's standard errors. Values significant at the 10%, 5% and 1% levels are marked by *, ** and ***, respectively.

	Bitfinex	**Bitstamp**	**BTC-e**	**ItBit**
		Volatility		
Mean	0.0221	0.0219	0.0213	0.0214
Minimum	0.0046	0.0041	0.0014	0.0000
Percentile 5	0.0063	0.0057	0.0054	0.0052
Median	0.0178	0.0177	0.0166	0.0163
Percentile 95	0.0509	0.0511	0.0559	0.0507
Maximum	0.0854	0.0900	0.0918	0.1484
Stand. deviation	0.0151	0.0155	0.0161	0.0174
Skewness	1.6374	1.8342	1.8138	2.9800
Kurtosis	5.9930	7.1510	6.7992	17.847
Jarque-Bera	164.02***	255.74***	229.94***	2133.0***
Autocorr(1)	0.4836***	0.4726***	0.4595***	0.2903***
Autocorr(2)	0.3057***	0.3057***	0.2563***	0.1559**
Autocorr(3)	0.3284***	0.3108***	0.2768***	0.2278***
BIC	1	1	1	1
		Volume		
Mean	9.6564	9.0602	8.7437	7.7818
Minimum	7.1639	7.6202	5.1789	0.0000
Percentile 5	8.2735	7.9872	7.9408	6.0695
Median	9.7017	9.0085	8.7168	7.8591
Percentile 95	11.016	10.206	9.5863	9.1890
Maximum	11.851	11.011	10.407	10.003
Stand. deviation	0.8555	0.7000	0.5836	1.0505
Skewness	-0.1557	0.1918	-1.0167	-2.1285
Kurtosis	2.7405	2.4600	9.7087	16.712
Jarque-Bera	1.3756	3.6745	411.56***	1726.4***
Autocorr(1)	0.7368***	0.7062***	0.4898***	0.5515***
Autocorr(2)	0.6095***	0.6230***	0.2259***	0.5079***
Autocorr(3)	0.5640***	0.6098***	0.2250**	0.4649***
BIC	1	1	1	2

CeBER Working Papers

(Available on-line at www.uc.pt/go/ceber *)*

2017-05 *Where is the information on USD/Bitcoins hourly price movements?*

2017-04 *The response of non-price competitiveness and productivity due to changes in passed income gaps. Evidence from the OECD countries - Pedro André Cerqueira, Micaela Antunes & Elias Soukiazis*

2017-03 *Dutch Disease in Central and Eastern European Countries - João Sousa Andrade & António Portugal Duarte*

2017-02 *On the gains of using high frequency data and higher moments in Portfolio Selection- Rui Pedro Brito, Hélder Sebastião & Pedro Godinho*

2017-01 *Growth adjustments through non-price competitiveness and productivity. A cumulative causation approach- Elias Soukiazis, Micaela Antunes & Pedro André Cerqueira*

A série CeBER Working Papers foi iniciada em 2017.