The dynamics of extreme price movements in cryptocurrencies

Christophe Desagre* Paolo Mazza† Mikael Petitjean‡

This version: July 27, 2019

Abstract

In this paper, we investigate the liquidity and trading dynamics around Extreme Price Movements (EPMs) in cryptocurrency markets. Based on Brogaard et al. (2018)'s methodology, we identify EPMs as periods when the absolute return is higher than a given threshold. We find that during EPMs, trading volume strongly increases, while the effect on liquidity is rather mixed as both depth and spreads increase. Next, we analyze whether these variables help explain the occurrence of an EPM. We start by analyzing Bitcoin in the platform Bitfinex. Then, we extend our investigation to a multi-platforms and a multi-cryptocurrencies analysis. In addition to a traditional LOGIT methodology, we also use FELOGIT and RELOGIT regressions. Our results suggest that trading activity, measured by the number of trades, and return are useful predictors to explain the occurrence of EPMs. As far as relative spread is concerned, the results are more mixed.

*Louvain School of Management (LFIN-IMMAQ, UCLouvain), 151 Chaussée de Binche, 7000 Mons, Belgium, Email : christophe.desagre@uclouvain.be. *Corresponding author.*

†IESEG School of Management, 3 rue de la Digue - 59000 Lille (France). E-mail: p.mazza@ieseg.fr.

‡Louvain School of Management (LFIN-IMMAQ, UCLouvain) and IESEG School of Management (LEM UMR 9221 CNRS). E-mail: m.petitjean@ieseg.fr.

I. Introduction

Over the last two decades, technological innovations have been particularly numerous in financial markets. One notable innovation is the development of a new type of currencies, known as cryptocurrencies. Cryptocurrencies allow for a digital form of payment in the sense that every transaction is recorded in a register called the *blockchain*. Among them, bitcoin (BTC) is the most famous. We witnessed an important liquidity dry-up during the last financial crisis. It is notoriously important to study liquidity, both in normal times and during crisis. As immature markets, cryptocurrencies deserve the same investigation. The growing popularity of cryptocurrencies have attracted the attention of academics, regulators, and central banks (Ali et al., 2014; McLeay et al., 2014).

The supply of these currencies is deterministically fixed. There is no central counter-party as it is the case with a central bank that chooses the quantity of money in circulation. Cryptocurrencies could completely reshape how financial systems work. However, the stability of this system is key to support its legitimacy. By solving the cryptographic problem, miners ensure the stability of the network. They are then rewarded with cryptocurrencies in exchange of their service. Cryptocurrencies markets operate 24/7. Therefore, there is no official closing prices, although we can use the last traded price of the day for empirical purposes. Importantly, cryptocurrencies can be traded against traditional currencies, e.g. USD, EUR, JPY, CNY, etc. or against other cryptocurrencies.

The introduction of the tracker by NASDAQ OMX in May 2015 and the two introductions of the futures by CBOE and CME in December 2017 are important milestones on Bitcoin's way to legitimacy as a financial asset. In that context, investors have to care not only about their potential capital gains, but also to the riskiness of their investment. Bitcoin can lose an important part of its value in a short time frame. For example, Stavroyiannis and Babalos (2017) report a 18% drop in Bitcoin's value in March 2017 because of SEC's denial to launch an ETF. Furthermore, Thies and Molnár (2018) report that daily returns vary from -48.52% to +40.14%. It is therefore of interest to better understand what triggers such extreme price movements.

In this paper, we address three research questions. Firstly, we look at what happens around extreme returns in cryptocurrencies markets in terms of liquidity and trading activity. Secondly, we analyze whether these crashes tend to recover. Finally, using a LOGIT regression approach, we investigate the potential drivers of these crashes. We start our analysis with BTCUSD on Bitfinex, and then we complement it by looking both at the cross-cryptocurrencies and the cross-platforms dynamics.

The remainder of the paper is as follows. In the next section, we review the relevant literature. Section III contains a description of our data and of our sample, the EPM identification strategy, the variables under scrutiny, some descriptive statistics, and the methodology. In Section IV, we report our empirical findings. Section V concludes.

II. Literature review

As far as cryptocurrencies are concerned, the financial academic literature is emerging. Still, we can separate most studies about Bitcoin, and cryptocurrencies in general, into 6 major research questions:

1. Are cryptocurrencies currencies?

2. Do cryptocurrencies have an intrinsic value?

3. Can we model cryptocurrencies' volatility?

4. What are the consequences of adding cryptocurrencies to a traditional portfolio?

5. Is there any price discovery across exchanges?

6. Can we explain cryptocurrencies' returns?

The first important question is to determine whether Bitcoin, and cryptocurrencies in general, are currencies, assets, or commodities. If we consider BTC as a currency, then its relationship with traditional currencies should be seen as an exchange rate. However, if we consider BTC as an asset or a commodity, then it has a price. This has implications when comparing cryptocurrencies to other asset classes. It is generally accepted that a money should have three functions, i.e. medium of exchange, store of value, and an unit of account (Ali et al., 2014). Glaser et al. (2014) document that Bitcoin users are rather interested in a new speculative instrument than in a system of payment. Rogojanu and Badea (2014) discuss the various attempts of alternative currencies through history and how Bitcoin relates to these other currencies. Dyhrberg (2016a, p. 85) considers BTC as *'something in between gold and the American Dollar on a scale from pure medium of exchange advantages to pure store of value advantages.'* Ali et al. (2014, p. 278) note that *'in contrast to commonly used forms of money such as banknotes or bank deposits, digital currencies are not a claim on anybody. In this respect, they can therefore be thought of as a type of commodity. But unlike physical commodities such as gold, they are also intangible assets, or digital commodities.'* It is notoriously important to understand the terminology as well as the difference between each concept. This question is important from a legislative point of view. Moreover, there exist a difference between Bitcoin (with a capital letter) which refers to the network and bitcoin (with a lower case) that refers to the cryptocurrency (Hale et al., 2018, p. 2). We report in Table A1 several definitions of bitcoin, cryptocurrency, virtual currency, etc. that we have identified in the literature.

Second, according to Hale et al. (2018, p. 1), Bitcoin is *'a cryptocurrency - a digital currency that is not backed by any tangible or intangible assets or intrinsic value.'* Some authors have questioned the intrinsic value of bitcoin. On the one hand, Cheah and Fry (2015) document that Bitcoin's fundamental value is zero as they find evidence of a bubble between January and November 2013. Donier and Bouchaud (2015, p. 2) note that *'the absence of any compelling way to assess the fundamental price of Bitcoins makes the behavioral hypothesis highly plausible'* to explain market crashes. This question is also closely related to the presence of a bubble in cryptocurrency markets.

Third, a stream of literature investigates how to model cryptocurrencies' volatility. Indeed, cryptocurrencies' volatility is high in comparison with other financial assets. Dwyer (2015) notes that Bitcoin exhibit higher volatility than currencies on average. Bouoiyour and Selmi (2015a) analyze Bitcoin price from December 2010 to June 2015. They use Threshold-GARCH (TGARCH) and Exponential GARCH (EGARCH). Bouoiyour et al. (2016) study Bitcoin volatility from December 2010 to July 2016. According to these authors, although volatility decreases in the second part of the sample, Bitcoin is still not a mature market. They use several GARCH-related models and discriminate among the models with information criteria. Dyhrberg (2016a) uses GARCH and EGARCH models. To date, there is no consensus on which GARCH model best works. The sample period under scrutiny is one possible explanation for these divergences of results.

Fourth, despite their high volatility, cryptocurrencies have been proposed as a new asset class to improve the risk-return trade-off in portfolio management. Cryptocurrencies are sometimes considered, rightly or wrongly, as the new gold (Dyhrberg, 2016a,b). According to Dyhrberg

(2016b), bitcoin exhibits low correlation with other asset classes, which makes it attractive for portfolio management. Briere et al. (2015) find that including Bitcoin in a diversified portfolio enhances the portfolio's performance. However, Bouoiyour and Selmi (2015b, p. 449) indicates that *"there is no sign of Bitcoin being a safe haven."*[1] On the contrary, the same authors in a different study mention that Bitcoin is a safe haven because of the system anonymity, which questions the relevance of their conclusions. Looking at the correlations across cryptocurrencies, Canh et al. (2019) show that the correlations are quite high, implying that it is difficult to diversify a portfolio composed only with cryptocurrencies.

Fifth, cryptocurrency markets are highly fragmented. In that context, some authors study whether a specific platform has an informational advantage in comparison to the others. Brandvold et al. (2015) analyze the price discovery across 7 platforms between April 2013 and February 2014 and find that MtGox and BTC-e (which had an important market share at that time, see Figure A2) drive Bitcoin price the most. Since MtGox shut down in the meantime, it is important to look at the cross-platforms dynamics. However, he notes that the information share strongly evolves over time.

Finally, a bunch of studies look whether there exists any predictability in cryptocurrencies returns. Kristoufek (2013) finds bidirectional relationships between bitcoin price and Internet metrics, i.e. Google Trends and Wikipedia. Garcia et al. (2014) also show that search attention (Google Trends) and the number of new bitcoin users help explain the variation of Bitcoin price. Baek and Elbeck (2015) and Liu et al. (2019) do not find any relationship between economic fundamentals and Bitcoin returns. They document that these returns are therefore a consequence

[1] According to Baur and Lucey (2010, p. 219), a safe haven is *"uncorrelated on the average with another asset, but negatively correlated during a market crisis."*

of participants' activity. Among 15 potential candidates, Chu et al. (2015) find that the generalized hyperbolic distribution to fit Bitcoin returns is the best statistical parametric distribution, while the normal distribution performs the worst. Financial assets exhibit non-normal returns. This fact has been widely documented in the literature. The same applies to cryptocurrencies and this non normality is even more pronounced. Using a Bayesian change point analysis, Thies and Molnár (2018) identify 48 structural breaks between September 2011 and August 2017 in Bitcoin returns.

To sum up, our study fits in the last research question, i.e. can we explain cryptocurrencies' returns? More specifically, we are interested in (i) analyzing what are the trading and liquidity dynamics around EPMs and (ii) predicting the occurrence of these events.

III. Empirical part

In this section, we first describe our databases, the sample of cryptocurrencies, and the sample of exchanges that we analyze in this paper. We list the variables used in the subsequent analysis. Then, we briefly mention our identification strategy which follows Brogaard et al. (2018)'s methodology.

A. Data and sample

We obtain data from Kaiko, an independent data provider that collects data directly from the exchanges. We have two datasets. In the first one, we have all trades that occured on the platforms with date and time, price, number of cryptocurrencies exchanged, and an indicator about whether the trade is buyer- or seller-initiated. In the second dataset, we have minutely orderbook snapshots with bid/ask price and quantities up to the 10^{th} limit.

The sample of platforms includes 16 exchanges on which BTCUSD is traded (i.e. Bitfinex, Bitflyer, Bitstamp, Bittrex, BTCC, BTCE, Cexio, Coinbase, Gatecoin, Gemini, Hitbtc, Huobi, Itbit, Kraken, OkCoin, and Quoine). Although the trade dataset starts in 2010, we only observe orderbook data since May 2015. Therefore, we restrict our analysis to the period May 2015 to July 2018 for which we have both trade and orderbook information. During this period, BTC experienced a sharpe increase of its price, reaching almost 20,000$ in December 2017, before declining to approximately 3,000$ a year later. In accordance with Liu et al. (2019), we divide our sample period into pre-, bubble, and post-bubble subsample periods.

In Table 1, we indicate the start and the end of the period for which we have orderbook information, the number of days, the number of observations, the daily average number of observations[2], the total number of trades, and the daily average of trades. As far as BTCUSD is concerned, Bitfinex is the biggest platform in terms of trading activity with more than 38,000 trades per day, followed by Coinbase, Hitbtc, Huobi, and Bitstamp. Over the period, we observe more than 180 million trades. In Panel B, we report for the platform Bitfinex the major cryptocurrencies traded against USD, i.e. Bitcoin Cash (BCH), Bitcoin (BTC), EOS (EOS), Ethereum (ETH), Litecoin (LTC), Stellar (XLM), Monero (XMR), and Ripple (XRP). In Panel C, we report for the platform Bitfinex BTC against traditional currencies, i.e. EUR, JPY, and USD.

As indicated in Table 1, cryptocurrencies platforms are numerous and this environment is highly competitive. Depending on the cryptocurrency, the number of platforms ranges from 1 to 16. In Figure A1, we represent the monthly market share of each platform for BTCUSD from May 2015

[2]We report this information as a proxy for data quality. This number should theoretically be equal to 1,440 as we should have one observation per minute. However, technical glitches or platforms' upgrades may affect this number. Indeed, one unique feature is that some platforms may temporarily be down because of hacking events or system upgrades. At the end of the sample period, the data provider changes its frequency (2,880 snapshots per day)

Table 1: Descriptive statistics

Panel A: BTCUSD

Exchange	Start	End	Nb days	Nb ob	Avg_ob	Nb trades	Avg_trades
bitfinex	15-May-15	21-Jul-18	1,163	1,537,507	1,322	45,133,393	38,808
bitflyer	18-Apr-18	21-Jul-18	94	239,226	2,545	109,869	1,169
bitstamp	15-May-15	20-Jul-18	1,162	1,521,183	1,309	19,608,226	16,875
bittrex	1-Jun-18	20-Jul-18	49	143,323	2,925	9,948	203
btcc	13-Feb-18	20-Jun-18	127	276,122	2,174	15,235	120
btce	15-May-15	21-Jul-18	1,163	1,397,452	1,202	16,288,069	14,005
cexio	11-Dec-17	20-Jul-18	221	442,639	2,003	2,885,221	13,055
coinbase	15-May-15	20-Jul-18	1,162	1,574,932	1,355	42,520,453	36,592
gatecoin	18-Feb-16	21-Jul-18	884	731,320	827	322,237	365
gemini	12-Oct-15	21-Jul-18	1,013	1,373,409	1,356	8,601,783	8,491
hitbtc	26-Aug-17	21-Jul-18	329	605,017	1,839	9,075,776	27,586
huobi	10-Nov-15	13-Sep-17	673	687,214	1,021	13,895,640	20,647
itbit	7-Oct-15	21-Jul-18	1,018	1,336,609	1,313	2,684,271	2,637
kraken	25-Aug-15	21-Jul-18	1,061	1,447,116	1,364	10,298,238	9,706
okcoin	15-May-15	21-Jul-18	1,163	1,566,039	1,347	10,403,684	8,946
quoine	22-Sep-16	21-Jul-18	667	948,031	1,421	2,127,941	3,190
TOTAL						**183,979,984**	

Panel B: Bitfinex - .../USD

Crypto	Start	End	Nb days	Nb ob	Avg_ob	Nb trades	Avg_trades
bchusd	10-Aug-17	21-Jul-18	345	615,299	1,783	9,982,619	28,935
btcusd	15-May-15	21-Jul-18	1,163	1,537,507	1,322	45,133,393	38,808
eosusd	10-Aug-17	21-Jul-18	345	636,848	1,846	12,757,782	36,979
ethusd	28-Apr-16	21-Jul-18	814	1,048,549	1,288	24,009,056	29,495
ltcusd	14-Sep-16	21-Jul-18	675	941,654	1,395	14,194,655	21,029
xlmusd	2-May-18	21-Jul-18	80	206,667	2,583	46,179	577
xmrusd	10-Aug-17	21-Jul-18	345	633,362	1,836	2,620,766	7,596
xrpusd	10-Aug-17	21-Jul-18	345	616,421	1,787	13,484,949	39,087
TOTAL						**123,823,971**	

Panel C: Bitfinex - BTC/...

Crypto	Start	End	Nb days	Nb ob	Avg_ob	Nb trades	Avg_trades
btceur	22-Nov-17	21-Jul-18	241	445,568	1,849	1,594,572	6,616
btcusd	15-May-15	21-Jul-18	1,163	1,537,507	1,322	45,133,393	38,808
btcjpy	29-Mar-18	21-Jul-18	114	253,095	2,220	110,254	967
TOTAL						**46,838,219**	

This Table reports the start and the end of the period for which we have orderbook information, the number of days, the number of observations, the daily average number of observations, the total number of trades, and the daily average of trades. In Panel A, we report all platforms on which BTCUSD is traded; in Panel B, we report all the cryptocurrencies traded in USD in Bitfinex; and in Panel C, we report the three fiat currencies against the BTC in Bitfinex.

Table 2: Analysis of empirical studies

Studies	Cryptocurrencies	Platform	Sample period
Kristoufek (2013)	Bitcoin	MtGox	May 2011 – June 2013
Garcia et al. (2014)	Bitcoin	MtGox	Jan. 2009 – Oct. 2013
Glaser et al. (2014)	Bitcoin	MtGox	Jan. 2011 – Oct. 2013
Baek and Elbeck (2015)	Bitcoin	www.bitcoincharts.com	July 2010 – Feb. 2014
Bouoiyour and Selmi (2015a)	Bitcoin	www.blockchain.info	Dec. 2010 – June 2015
Bouoiyour and Selmi (2015b)	Bitcoin	www.blockchain.info	Dec. 2010 – June 2014
Brandvold et al. (2015)	Bitcoin	7 platforms	Apr. 2013 – Feb. 2014
Briere et al. (2015)	Bitcoin	www.bitcoincharts.com	July 2010 – Dec. 2013
Cheah and Fry (2015)	Bitcoin	Bitcoin Coindesk	July 2010 – July 2014
Chu et al. (2015)	Bitcoin	Bitstamp	Sep. 2011 – May 2014
Donier and Bouchaud (2015)	Bitcoin	MtGox	Dec. 2011 – Jan. 2014
Dwyer (2015)	Bitcoin	3 platforms	July 2010 – Apr. 2014
Bouoiyour et al. (2016)	Bitcoin	www.blockchain.info	Dec. 2010 – July 2016
Dyhrberg (2016b)	Bitcoin	Coindesk BPI	July 2010 – May 2015
Ardia et al. (2018)	Bitcoin	Datastream	Aug. 2011 – Mar. 2018
Feng et al. (2018)	Bitcoin	Bitstamp	Sep. 2011 – July 2017
Thies and Molnár (2018)	Bitcoin	Bitstamp	Sep. 2011 – Aug. 2017
Canh et al. (2019)	7 cryptocurrencies	Coinmarketcap	Aug. 2014 – Dec. 2018

This Table reports the cryptocurrencies, platforms, and time window under consideration for the studies mentioned in our literature review.

to July 2018[3].

Some empirical studies only use daily data (i.e. Open-High-Low-Close prices and volume information). These data do not bring any information on the intraday price dynamics. In Table 2, we report for each empirical study the cryptocurrencies, platforms[4], and time window under consideration.

[3]In Appendix - Figure A2, we report the same figure from July 2010 to September 2018. We observe that in the beginning of the period, there was a monopolistic situation held by MtGox. However, this platform shut down in February 2014, which resulted in a loss of more than 400 millions of dollars for its users according to Forbes (Source: `https://www.forbes.com/sites/cameronkeng/2014/02/25/bitcoins-mt-gox-shuts-down-loses-409200000-dollars-recovery-steps-and-taking-your-tax-losses/#41ba609d5c16`).

[4]When platform is not mentioned, we report the website from which data are retrieved. These website provide an average of Bitcoin/cryptocurrencies price across platforms, sometimes weighted by the importance of the platform.

B. Variables

From the orderbook dataset, we compute quoted spreads (QS), relative spreads (RS), depth at the best quotes ($DEPTH$) and at the 5 best quotes ($DEPTH5$) in monetary volume to proxy for liquidity. We measure the orderbook imbalance at the best quotes (OB_Imb) and at the 5 best quotes (OB_Imb5) as:

$$OB\_Imb = \frac{QB_1 - QA_1}{QB_1 + QA_1} \tag{1}$$

$$OB\_Imb5 = \frac{\sum_{i=1}^{5}(QB_i - QA_i)}{\sum_{i=1}^{5}(QB_i + QA_i)} \tag{2}$$

with QB_i (QA_i) the quantity available at limit i. In accordance with Brogaard et al. (2018), we compute returns (R_t) using the midpoint. For each hourly interval, we compute the average and median values of these variables.

From the trade dataset, we measure the number of trades (NT), the quantities traded (QT), the volume traded (VT), the average trade size ($ATS = QT/NT$), and the average trade volume ($ATV = VT/NT$)[5]. We compute Amihud (2002)'s measure:

$$Amihud_t = \frac{|R_t|}{VT_t} \tag{3}$$

Large price movements can be triggered by at least two types of events: information arrival and

[5] As indicated in Figure A3, both variables are strongly correlated with BTCUSD price. Over the sample period, we measure a correlation of -0.52 (0.81) between ATS (ATV) and price.

trade imbalances (Brogaard et al., 2018, p. 258). We measure trade imbalance (T_Imb_t) as:

$$T\_Imb_t = \frac{SELL_t - BUY_t}{SELL_t + BUY_t} \tag{4}$$

where BUY_t ($SELL_t$) is the number of buyer-initiated (seller-initiated) trades during interval t. We do not require to use Lee and Ready (1991)'s algorithm as this information is provided in the database[6]. We report descriptive statistics in Table 3 (Panel A).

C. Methodology

This methodological section is divided into two main parts. First, we explain our identification strategy, which is similar to Brogaard et al. (2018)'s methodology. Second, we discuss the potential alternative to estimate a LOGIT regression, taking into account the fixed effects potentially present in our data and the scarcity of EPMs.

Brogaard et al. (2018, p. 253) define *stressful periods as 'unexpected and rapidly developing extreme price movements (EPMs) that belong to the* 99.9^{th} *percentile of the return distribution'*. They also define co-EPMs as *'an instance where more than one stock simultaneously undergoes an EPM.'* We follow this approach to identify EPMs by computing absolute logarithmic returns based on the last observed midpoint during the interval. While these authors use a 10-second interval for stocks traded on NASDAQ, we decide to take a longer interval, i.e. one hour, as cryptocurrencies markets are relatively immature in comparison with equity markets. In robustness checks, we replicate our analysis for two other interval lengths, i.e. 15 minutes and 30 minutes.

[6]For BTCUSD, this information is available from April, 2 2014. This is not an issue since we focus on a subsample period from May 2015 to July 2018

To increase the number of EPMs in our sample, we also use the threshold of 99^{th} percentile. Brogaard et al. (2018) analyze two years of data and use a frequency of 10 second, resulting in more than 45 millions observations, and 45,200 EPMs. As we work with an hourly interval from May 2015 to July 2018, we only observe around 25 EPMs if we use a threshold the 99.9 percentile. Allowing the threshold to be 99^{th} increases our sample of EPMs to 250. From now, we will refer to returns exceeding the 99^{th} (99.9^{th}) percentile as EPM_{99} ($EPM_{99.9}$).

As a first step, we use a logistic (LOGIT) regression framework to identify whether any of the aforementioned variables can predict the occurrence of an EPM. Formally, we specify our LOGIT model as follows:[7]

$$Prob(y_t = 1|\boldsymbol{x_t'}) = \frac{exp(\boldsymbol{x_t'\beta})}{1 + exp(\boldsymbol{x_t'\beta})} \tag{5}$$

where y_t is the EPM variable; $\boldsymbol{x_t'}$ is a 1x(k+1) vector of the k explanatory variables, including the intercept and four variables indexed at time $t-1$, i.e. return, imbalance, number of trades, and RS; and $\boldsymbol{\beta}$ is a (k+1)x1 vector of coefficients. All non-dummy variables are standardized. In comparison with Brogaard et al. (2018), we use the number of trades, and not the share volume. This choice is motivated by the fact that the number of cryptocurrencies traded is strongly impacted by its price (as we showed in Figure A3) and that there exist important disparities across cryptocurrencies with respect to their price. As we can trade fractions of cryptocurrencies, the number of trades is less impacted by the cryptocurrency's price.

By definition, there is one EPM for 999/99 non-EPM in our baseline model. Consequently, these events are extremely rare and the use of a LOGIT regression may induce some biases due to the

[7]Notations are taken from Mazza (2019).

disequilibrium between the number of events ($EPM_{99/99.9} = 1$) and non-events ($EPM_{99/99.9} = 0$). To take this imbalance into account, we extend our baseline model to a a Rare Event LOGIT (RELOGIT). This method is discussed by King and Zeng (2001a,b) and by Cook et al. (2018) in the field of political science and international conflicts. Mazza (2019) compares several logistic models and indicates that a LOGIT with fixed effects (FELOGIT) and conditional logit (CLOGIT) are the best alternatives when it comes to analyze rare events, while controlling for fixed effects.

Then, we extend our analysis of BTCUSD on Bitfinex to other cryptocurrencies and other platforms. Given that, we decide to include cryptocurrencies- and platforms-fixed effects in the regression, which results in a panel analysis. Including platforms fixed effects, cryptocurrencies fixed effects, or both effects conjointly, Equation 5 becomes respectively:

$$Prob(y_{it} = 1 | \boldsymbol{x'_{it}}, \alpha_i) = \frac{exp(\boldsymbol{x'_{it}}\boldsymbol{\beta} + \alpha_i)}{1 + exp(\boldsymbol{x'_{it}}\boldsymbol{\beta} + \alpha_i)} \tag{6}$$

$$Prob(y_{jt} = 1 | \boldsymbol{x'_{jt}}, \alpha_j) = \frac{exp(\boldsymbol{x'_{jt}}\boldsymbol{\beta} + \alpha_j)}{1 + exp(\boldsymbol{x'_{jt}}\boldsymbol{\beta} + \alpha_j)} \tag{7}$$

$$Prob(y_{ijt} = 1 | \boldsymbol{x'_{ijt}}, \alpha_i, \alpha_j) = \frac{exp(\boldsymbol{x'_{ijt}}\boldsymbol{\beta} + \alpha_i + \alpha_j)}{1 + exp(\boldsymbol{x'_{ijt}}\boldsymbol{\beta} + \alpha_i + \alpha_j)} \tag{8}$$

where i (j) is a subscript for platform (cryptocurrency), and α_i (α_j) represents a platform- (cryptocurrency)-fixed effect. Given that the model is non-linear, fixed effects do not disappear from the likelihood function and we need to correct for it. However, as the number of intervals, t, is large and the number of individuals i (j), i.e. the number of platforms (cryptocurrencies) is small, the bias should not be important.

IV. Empirical results

A. Descriptive statistics

In Table A2, we report different percentiles (i.e. 50 ; 75 ; 90 ; 95 ; 99 ; 99.9 ; 99.99) for absolute log-returns computed for an interval of 3600s for platform Bitfinex. We compute returns based on the last midpoint of the interval[8]. The returns are expressed in percentage. For our main analysis (Bitfinex-BTCUSD), there is an EPM_{99} ($EPM_{99.9}$) when the absolute return during the interval exceeds 11.58% (25.97%).

We identify 275 (27) EPM_{99} ($EPM_{99.9}$). As shown in Figure A4, we are not able to distinguish any clear intraday pattern for the occurrence of EPMs. For each EPM, we analyze its percentage of recovery up to 24 hours after the EPM. We report in Figure A6 how much do these EPMs recover. We distinguish between "down" EPMs and "up" EPMs.

After identifying these EPMs, we look at the liquidity and trading dynamics around them. To illustrate our point, Figure A5 shows the liquidity and trading dynamics during an EPM, there is also a spike in volume (above graph). At the same time, we show that the relative spread increases significantly during an EPM (below graph). Both findings suggest conflicting results as higher trading activity (higher spreads) is positive (negative) in terms of liquidity.

In Tables 3 and 4, we report the mean and median values of average trade size (ATS), average trade volume (ATV), depth at best quotes ($Depth$), depth at the 5 best quotes ($Depth5$), number of trades (NT), orderbook imbalance (OB_Imb), orderbook imbalance at the 5 best quotes

[8]In case of missing data during an interval, we do not take into account the return to avoid computing returns over several intervals.

(OB_Imb5), relative spread (RS), quoted spread (QS), quantities traded (QT), trade imbalance (T_Imb), and volume traded (VT) estimated during both the full sample and during EPMs. Consistent with Figure A5, we find that the average relative spread during EPMs is almost twice its value in comparison with the full sample. Trading activity, be it measured in number of trades, or ATV also strongly increase during EPMs. The median values confirm these findings. We test for the statistical significance of these differences. Table 3 (4) compares the full sample and EPM_{99} ($EPM_{99.9}$). Interestingly, the differences between the full sample and $EPM_{99.9}$ are more important than the differences between the full sample and EPM_{99}, which makes economic sense. In appendix, we replicate this analysis for intervals of 30 minutes (Table A3) and 15 minutes (Table A4).

B. BTCUSD - Bitfinex

In Table 5, we report the results of Equation 5. We estimate the model both for EPM_{99} and $EPM_{99.9}$. The dependent variable is the occurrence of an EPM at time t and the independent variables include an intercept, the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS). All variables, excepting the intercept, are lagged by one period and are standardized. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. In Panel A, we estimate a LOGIT regression. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. According to Mazza (2019, p. 8), this correction is *'a perfect solution for rare events.'* In Panel C, we estimate a LOGIT regression with Firth (1993)'s correction and where confidence intervals are computed based on the profile penalized log likelihood.

Table 3: Descriptive statistics

	Mean	Median	StDev.	N
Panel A: Full sample				
ATS	1.37	0.94	1.19	27,505
ATV	1,679.40	1,229.77	1,295.42	27,505
$Depth$	34,802.48	14,196.59	54,852.16	25,295
$Depth5$	108,006.40	57,436.09	120,752.60	25,295
NT	1,603.36	513.00	2,599.95	27,909
OB_Imb	(2.22)	(0.79)	8.44	25,295
OB_Imb5	(0.00)	0.00	0.24	25,295
RS	0.04	0.02	0.19	25,295
QS	0.63	0.23	1.52	25,295
QT	1,125.28	584.58	1,641.52	27,909
T_Imb	(0.02)	(0.02)	0.29	27,909
VT	5,125,542.00	522,850.50	11,433,054.00	27,909
Panel B: Extreme price movements (EPM_{99})				
ATS	1.26	0.56	1.41	275
ATV	3,350.16***	3,539.53	1,520.06	275
$Depth$	60,881.66***	51,103.14	62,855.32	251
$Depth5$	215,201.80***	192,957.30	179,184.00	251
NT	9,312.30 ***	9,214.00	6,796.36	275
OB_Imb	(2.62)	(1.36)	4.02	251
OB_Imb5	0.03 ***	0.02	0.16	251
RS	0.07***	0.05	0.08	251
QS	3.24***	2.40	3.06	251
QT	7,184.26***	5,644.42	6,551.86	275
T_Imb	(0.05)***	(0.06)	0.17	275
VT	38,008,423.00***	35,532,461.00	33,612,221.00	275

This Table reports descriptive statistics about our variables: average trade size (ATS), average trade volume (ATV), depth at best quotes ($Depth$), depth at the 5 best quotes ($Depth5$), number of trades (NT), orderbook imbalance (OB_Imb), orderbook imbalance at the 5 best quotes (OB_Imb5), relative spread (RS), quoted spread (QS), quantities traded (QT), trade imbalance (T_Imb), and volume traded (VT). All variables are defined in section B. For each variable, we report the mean, median, standard deviation, and number of observations. We test for the statistical differences between Panel A and Panel B. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table 4: Descriptive statistics

	Mean	Median	StDev.	N
Panel A: Full sample				
ATS	1.37	0.94	1.19	27,505
ATV	1,679.40	1,229.77	1,295.42	27,505
$Depth$	34,802.48	14,196.59	54,852.16	25,295
$Depth5$	108,006.40	57,436.09	120,752.60	25,295
NT	1,603.36	513.00	2,599.95	27,909
OB_Imb	(2.22)	(0.79)	8.44	25,295
OB_Imb5	(0.00)	0.00	0.24	25,295
RS	0.04	0.02	0.19	25,295
QS	0.63	0.23	1.52	25,295
QT	1,125.28	584.58	1,641.52	27,909
T_Imb	(0.02)	(0.02)	0.29	27,909
VT	5,125,542.00	522,850.50	11,433,054.00	27,909
Panel C: Extreme price movements ($EPM_{99.9}$)				
ATS	1.85	0.62	2.02	27
ATV	3,616.06***	3,710.30	1,877.23	27
$Depth$	65,165.39***	60,015.48	55,183.23	25
$Depth5$	242,795.90***	198,610.00	190,966.30	25
NT	14,267.22***	15,483.00	9,795.76	27
OB_Imb	(3.06)	(2.07)	4.20	25
OB_Imb5	0.02	(0.00)	0.17	25
RS	0.15***	0.09	0.17	25
QS	4.74***	3.02	4.68	25
QT	15,108.05***	11,005.31	11,645.53	27
T_Imb	-0.0714	-0.07515	0.16739	27
VT	62,064,214.00***	69,237,964.00	50,515,370.00	27

This Table reports descriptive statistics about our variables: average trade size (ATS), average trade volume (ATV), depth at best quotes ($Depth$), depth at the 5 best quotes ($Depth5$), number of trades (NT), orderbook imbalance (OB_Imb), orderbook imbalance at the 5 best quotes (OB_Imb5), relative spread (RS), quoted spread (QS), quantities traded (QT), trade imbalance (T_Imb), and volume traded (VT). All variables are defined in section B. For each variable, we report the mean, median, standard deviation, and number of observations. We test for the statistical differences between Panel A and Panel C. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

For Panel A and EPM_{99}, we find that two variables are statistically significant, i.e. the number of trades and the return. Both variables exhibit a positive coefficient. Looking at $EPM_{99.9}$, we find consistent results. When using Firth (1993)'s correction (Panel B), we observe that the relative spread becomes statistically significant. This variable also exhibits a positive coefficient. Results are consistent when we correct using Firth (1993)'s correction and when confidence intervals are computed based on the profile penalized log likelihood (Panel C). In appendix, we replicate this analysis when variables are averaged over 30 minutes (Tables A9).

To take the imbalance between events and non-events, we also estimate a RELOGIT. We report results in Table 6. All reported coefficients are in accordance with those previously documented.

C. Cross-platforms analysis

While Bitfinex is the most important platform in terms of trading activity, it represents less than 25% of the trades for BTCUSD. As indicated in Table 1, there are 15 other platforms on which investors can trade BTCUSD. To take these other platforms into account in our analysis, we estimate Equation 6. In comparison with Equation 5, we include a fixed effect for each platform (and consequently, we remove the intercept). All non-dummy variables are standardized at the platform-level. We report the estimation results in Table 7[9].

We find that the three variables identified as statistically significant in Table 5 are still statistically significant and the sign of the coefficient is consistent. Trade imbalance also becomes a statistically significant variable at the 1% level for EPM_{99}, but not for $EPM_{99.9}$. Platforms-fixed effects considerably improve the quality of the model.

[9]In Appendix, we report the results with the values of platforms' fixed effects, see Tables A5 and A6

Table 5: LOGIT - Bitfinex - BTCUSD

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
α_0	(4.9951)	***	(7.2739)	***
NT_{t-1}	0.4741	***	0.4214	***
T_Imb_{t-1}	(0.0448)		(0.0353)	
R_{t-1}	0.191	***	0.1741	**
RS_{t-1}	0.0294		0.0361	
N	25,175		25,175	
$N_{y=0}$	24,925	99.01%	25,149	99.90%
$N_{y=1}$	250	0.99%	26	0.10%
R^2	1.49%		0.19%	
Panel B: LOGIT with FIRTH correction				
	EPM_{99}		$EPM_{99.9}$	
α_0	(4.9881)	***	(7.2145)	***
NT_{t-1}	0.4738	***	0.4131	***
T_Imb_{t-1}	(0.0451)		(0.0364)	
R_{t-1}	0.1907	***	0.1878	**
RS_{t-1}	0.0357	**	0.0517	***
N	25,175		25,175	
$N_{y=0}$	24,925	99.01%	25,149	99.90%
$N_{y=1}$	250	0.99%	26	0.10%
R^2	1.51%		0.22%	
Panel C: LOGIT with FIRTH correction and penalized log likelihood				
	EPM_{99}		$EPM_{99.9}$	
α_0	(4.9881)	***	(7.2145)	***
NT_{t-1}	0.4738	***	0.4131	***
T_Imb_{t-1}	(0.0451)		(0.0364)	
R_{t-1}	0.1907	***	0.1878	**
RS_{t-1}	0.0358	**	0.0517	**
N	25,175		25,175	
$N_{y=0}$	24,925	99.01%	25,149	99.90%
$N_{y=1}$	250	0.99%	26	0.10%
R^2	1.51%		0.22%	

This Table reports results of Equation 5. The dependent variable is the occurrence of an EPM at time t and the independent variables include an intercept, the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS). All variables, excepting the intercept, are lagged by one period and are standardized. In Panel A, we estimate a LOGIT regression. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. In Panel C, we estimate a LOGIT regression with Firth (1993)'s correction and where confidence intervals are computed based on the profile penalized log likelihood. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table 6: RELOGIT - Bitfinex - BTCUSD

	EPM_{99}		$EPM_{99.9}$	
α_0	(4,9809)	***	(7,2244)	***
NT_{t-1}	0,4740	***	0,4153	***
T_Imb_{t-1}	(0,0450)		(0,0357)	
R_{t-1}	0,1902	***	0,185	*
RS_{t-1}	0,0450	***	0,1054	***
N	25,175		25,175	
$N_{y=0}$	24,925	99.01%	25,149	99.90%
$N_{y=1}$	250	0.99%	26	0.10%

This Table reports results of Equation 5. The dependent variable is the occurrence of an EPM at time t and the independent variables include an intercept, the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS). All variables, excepting the intercept, are lagged by one period and are standardized. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

D. Cross-cryptocurrencies analysis

In Figure A7, we provide graphical evidence that during an EPM, there is a close relationship between the price of BTCUSD and ETHUSD. The above graph shows that both prices tend to move together. The below graph indicates that trading activity in both cryptocurrencies is strongly correlated. To take the cross-cryptocurrencies dynamics into account, we estimate Equation 7. In comparison with Equation 5, we include cryptocurrencies-fixed effects, and we remove the intercept as in Equation 6. To control for other effects potentially related to the platform, we look at the cross-cryptocurrencies dynamics occurring in Bitfinex. All non-dummy variables are standardized at the cryptocurrency-level. Results are in Table 8[10].

We find that the return remains statistically significant at the 1% level. When comparing EPM_{99} and $EPM_{99.9}$, we find mixed results for the number of trades and the relative spread. Both variables are significant for EPM_{99} and not for $EPM_{99.9}$. Results are qualitatively similar when using Firth (1993)'s correction.

[10] In Appendix, we report the results with the values of cryptocurrencies' fixed effects, see Tables A7

Table 7: LOGIT - BTCUSD - All exchanges

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
NT_{t-1}	0.2429	***	0.1902	***
T_Imb_{t-1}	(0.0994)	***	(0.1157)	
R_{t-1}	0.3729	***	0.3158	***
RS_{t-1}	0.1275	***	0.0526	***
α_i	YES		YES	
N	222,009		222,009	
$N_{y=0}$	219,799	99.00%	221,788	99.90%
$N_{y=1}$	2,210	1.00%	221	0.10%
R^2	72.42%		74.65%	
Panel B: LOGIT with FIRTH correction				
	EPM_{99}		$EPM_{99.9}$	
NT_{t-1}	0.2427	***	0.1896	***
T_Imb_{t-1}	(0.0994)	***	(0.1159)	*
R_{t-1}	0.3728	***	0.3156	***
RS_{t-1}	0.1268	***	0.0531	***
α_i	YES		YES	
N	222,009		222,009	
$N_{y=0}$	219,799	99.00%	221,788	99.90%
$N_{y=1}$	2,210	1.00%	221	0.10%
R^2	72.41%		74.64%	

This Table reports results of Equation 6. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), and platforms' fixed effects. All non-dummy variables are standardized at the platform-level. In Panel A, we estimate a LOGIT regression. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table 8: LOGIT - Bitfinex - All cryptocurrencies

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.27	***	0.04	
T_Imb_{t-1}	(0.03)		(0.01)	
R_{t-1}	0.24	***	0.11	***
RS_{t-1}	0.04	***	(0.02)	
α_j	YES		YES	
N	98,176		98,176	
N_0	97,160	98.97%	98,084	99.91%
N_1	1,016	1.03%	92	0.09%
R2	70.09%		72.53%	
Panel B: LOGIT with FIRTH correction				
	EPM_{99}		$EPM_{99.9}$	
NT_{t-1}	0.27	***	0.04	
T_Imb_{t-1}	(0.03)		(0.01)	
R_{t-1}	0.24	***	0.11	***
RS_{t-1}	0.04	***	(0.02)	
α_j	YES		YES	
N	98,176		98,176	
$N_{y=0}$	97,160	98.97%	98,084	99.91%
$N_{y=1}$	1,016	1.03%	92	0.09%
R^2	70.08%		72.52%	

This Table reports results of Equation 7. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), and cryptocurrencies' fixed effects. All non-dummy variables are standardized at the cryptocurrency-level. In Panel A, we estimate a LOGIT regression. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

E. Cross platforms-cryptocurrencies dynamics

In section B, we analyze BTCUSD in Bitfinex. Then in sections C and D, we include respectively platforms- and cryptocurrencies-fixed effects. In this section, we analyze both fixed effects conjointly and estimate Equation 6. Results are reported in Table 9. In appendix, we report the values of these fixed effects (Table A8).

Again, we find that the variable NT display a positive and statistically significant coefficient (at the 1% level). Trading imbalance exhibits a negative coefficient, but it is only significant for EPM_{99}. Return and relative spreads display positive and statistically significant coefficients. In Table A8, we find that only platforms-fixed effects are significant. Results with Firth (1993)'s correction are qualitatively similar[11].

Table 9: LOGIT - All platforms - All cryptocurrencies

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.2723	***	0.2141	***
T_Imb_{t-1}	(0.0831)	***	(0.1047)	
R_{t-1}	0.3436	***	0.3093	***
RS_{t-1}	0.1609	***	0.0577	***
α_i	YES		YES	
α_j	YES		YES	
N	320,288		320,288	
$N_{y=0}$	317,055	98.99%	319,975	99.90%
$N_{y=1}$	3,233	1.01%	313	0.10%
R^2	72.38%		74.65%	

This Table reports results of Equation 7. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), platforms- and cryptocurrencies' fixed effects. All non-dummy variables are standardized at the platform-cryptocurrency-level. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

[11] These results are available upon request.

V. Conclusion

In this paper, we investigate the liquidity and trading dynamics around EPMs. Based on Brogaard et al. (2018)'s methodology, we identify EPMs as periods when the absolute return is higher than a given threshold. We find that during EPMs, trading volume strongly increases, while the effect on liquidity is rather mixed as both depth and spreads increase. Next, we analyze whether these variables help explain the occurrence of an EPM. We start by analyzing Bitcoin in the platform Bitfinex. Then, we extend our analysis to a multi-platforms and multi-cryptocurrencies universe. In addition to a traditional LOGIT methodology, we also use FELOGIT and RELOGIT regressions. Our results suggest that trading activity, measured by the number of trades and return are useful predictors to explain the occurrence of EPMs. As far as relative spread is concerned, the results are more mixed.

References

Ali, R., J. Barrdear, R. Clews, and J. Southgate (2014). The economics of digital currencies. *Quarterly Bulletin Q3*, 276–286.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Ardia, D., K. Bluteau, and M. Rüede (2018). Regime changes in bitcoin GARCH volatility dynamics. *Finance Research Letters 29*, 266–271.

Baek, C. and M. Elbeck (2015). Bitcoins as an investment or speculative vehicle? A first look. *Applied Economics Letters 22*(1), 30–34.

Balcilar, M., E. Bouri, R. Gupta, and D. Roubaud (2017). Can volume predict bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling 64*, 74–81.

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

Baur, D. G. and B. M. Lucey (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review 45*(2), 217–229.

Bouoiyour, J. and R. Selmi (2015a). Bitcoin price: is it really that new round of volatility can be on way? *Munich Personal RePEc Archive*, 1–14.

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

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

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

Briere, M., K. Oosterlinck, and A. Szafarz (2015). Virtual currency, tangible return: portfolio diversification with bitcoin. *Journal of Asset Management 16*(6), 365–373.

Brogaard, J., A. Carrion, T. Moyaert, R. Riordan, A. Shkilko, and K. Sokolov (2018). High frequency trading and extreme price movements. *Journal of Financial Economics 128*(2), 253–265.

Canh, N. P., U. Wongchoti, S. D. Thanh, and N. T. Thong (2019). Systematic risk in cryptocurrency market: Evidence from dcc-mgarch model. *Finance Research Letters 29*, 90–100.

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

Chu, J., S. Chan, S. Nadarajah, and J. Osterrieder (2017). GARCH modelling of cryptocurrencies. *Journal of Risk and Financial Management 10*(4), 1–15.

Chu, J., S. Nadarajah, and S. Chan (2015). Statistical analysis of the exchange rate of bitcoin. *PloS one 10*(7), e0133678.

Cook, S. J., J. C. Hays, and R. J. Franzese (2018). Fixed effects in rare events data: a penalized maximum likelihood solution. *Political Science Research and Methods*, 1–14.

Corbet, S., B. Lucey, A. Urquhart, and L. Yarovaya (2019). Cryptocurrencies as a financial asset: a systematic analysis. *International Review of Financial Analysis 62*, 182–199.

Donier, J. and J.-P. Bouchaud (2015). Why do markets crash? Bitcoin data offers unprecedented insights. *PloS one 10*(10), e0139356.

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

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

Dyhrberg, A. H. (2016b). Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters 16*, 139–144.

European Central Bank (2012). Virtual currency schemes. Technical report.

European Central Bank (2015). Virtual currency schemes–a further analysis.

Feng, W., Y. Wang, and Z. Zhang (2018). Informed trading in the bitcoin market. *Finance Research Letters 26*, 63–70.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika 80*(1), 27–38.

Garcia, D., C. J. Tessone, P. Mavrodiev, and N. Perony (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface 11*(99), 20140623.

Glaser, F., K. Zimmermann, M. Haferkorn, M. C. Weber, and M. Siering (2014). Bitcoin-asset or currency? Revealing users' hidden intentions.

Hale, G., A. Krishnamurthy, M. Kudlyak, and P. Shultz (2018). How futures trading changed bitcoin prices. *FRBSF Economic Letter 12*, 1–5.

Hayes, A. S. (2017). Cryptocurrency value formation: an empirical study leading to a cost of production model for valuing bitcoin. *Telematics and Informatics 34*(7), 1308–1321.

King, G. and L. Zeng (2001a). Explaining rare events in international relations. *International Organization 55*(3), 693–715.

King, G. and L. Zeng (2001b). Logistic regression in rare events data. *Political analysis 9*(2), 137–163.

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

Lee, C. M. and M. J. Ready (1991). Inferring trade direction from intraday data. *Journal of Finance 46*(2), 733–746.

Liu, J., I. W. Marsh, P. Mazza, and M. Petitjean (2019). Factor structure in cryptocurrency returns and volatility. *working paper*.

Mazza, P. (2019). Controlling for rare events and fixed effects in a logistic regression framework: A review of the alternatives. *working paper*.

McLeay, M., A. Radia, and R. Thomas (2014). Money creation in the modern economy. *Bank of England Quarterly Bulletin Q1*.

Osterrieder, J. and J. Lorenz (2017). A statistical risk assessment of bitcoin and its extreme tail behavior. *Annals of Financial Economics 12*(01), 1750003.

Rogojanu, A. and L. Badea (2014). The issue of competing currencies. Case study-bitcoin. *Theoretical & Applied Economics 21*(1).

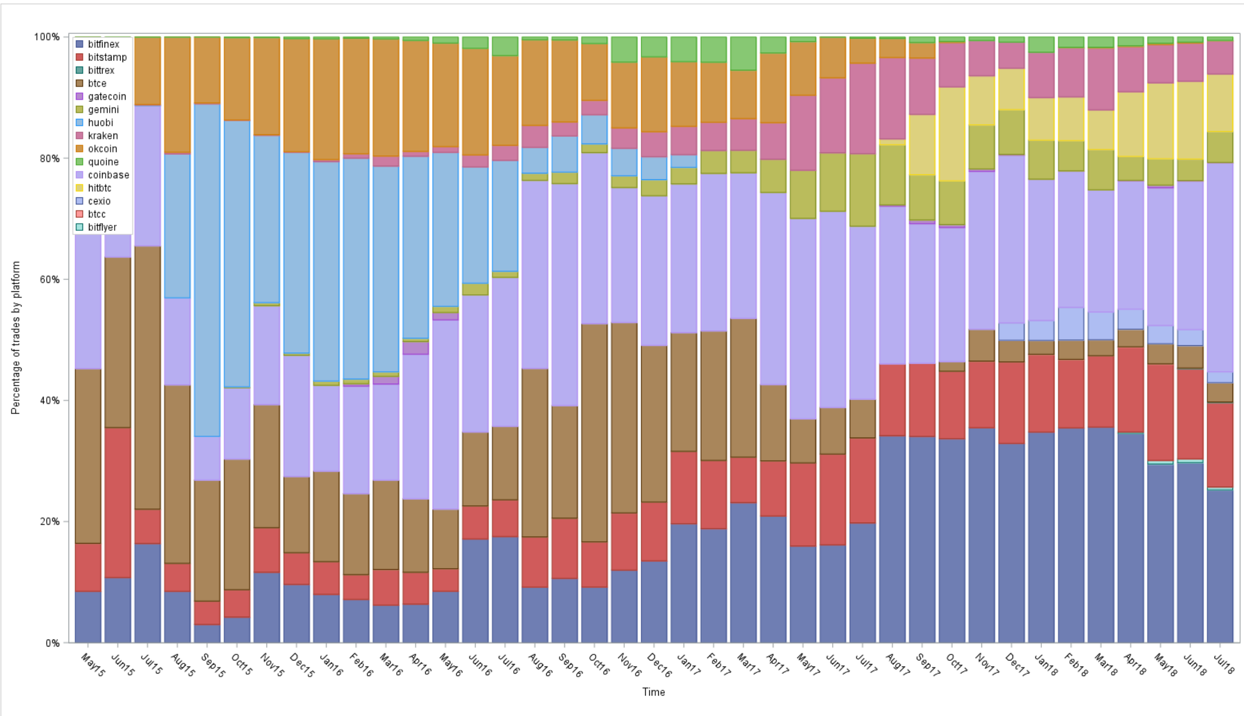
Stavroyiannis, S. and V. Babalos (2017). Dynamic properties of the bitcoin and the US market. *working paper*.

Thies, S. and P. Molnár (2018). Bayesian change point analysis of bitcoin returns. *Finance Research Letters 27*, 223–227.

VI. Appendix

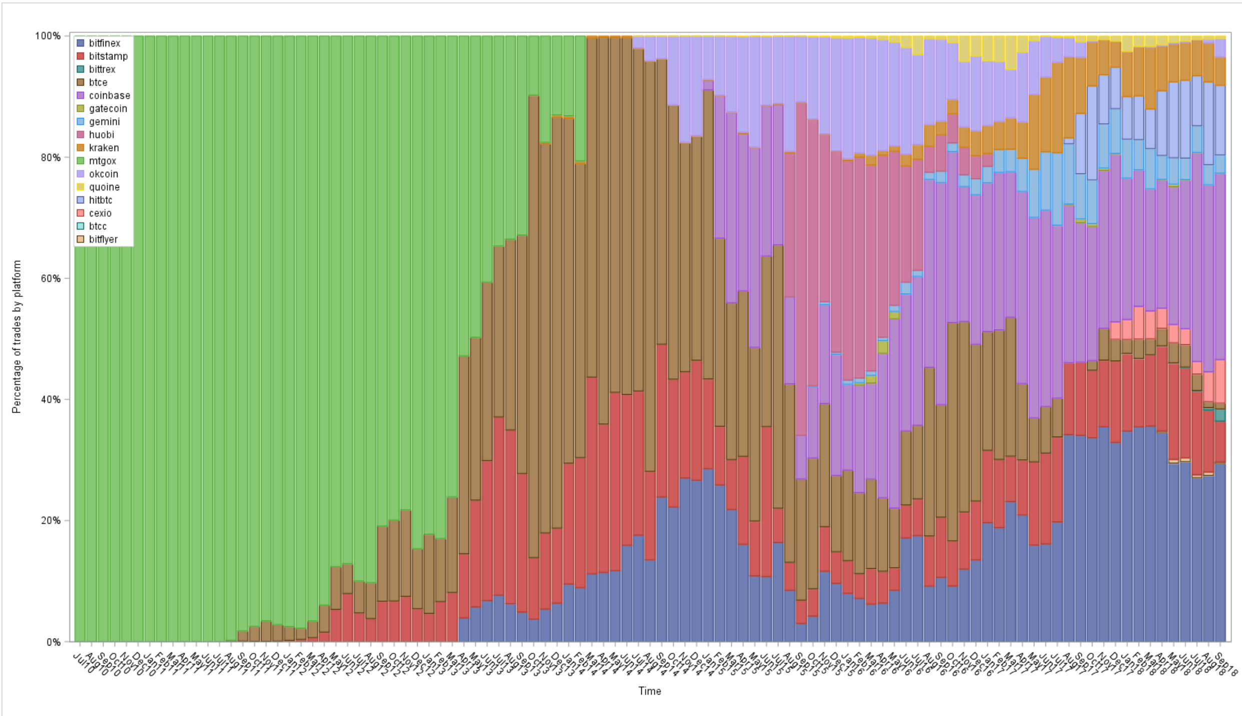
A. Figures

Figure A1: Market share across platforms (May 2015 to July 2018)



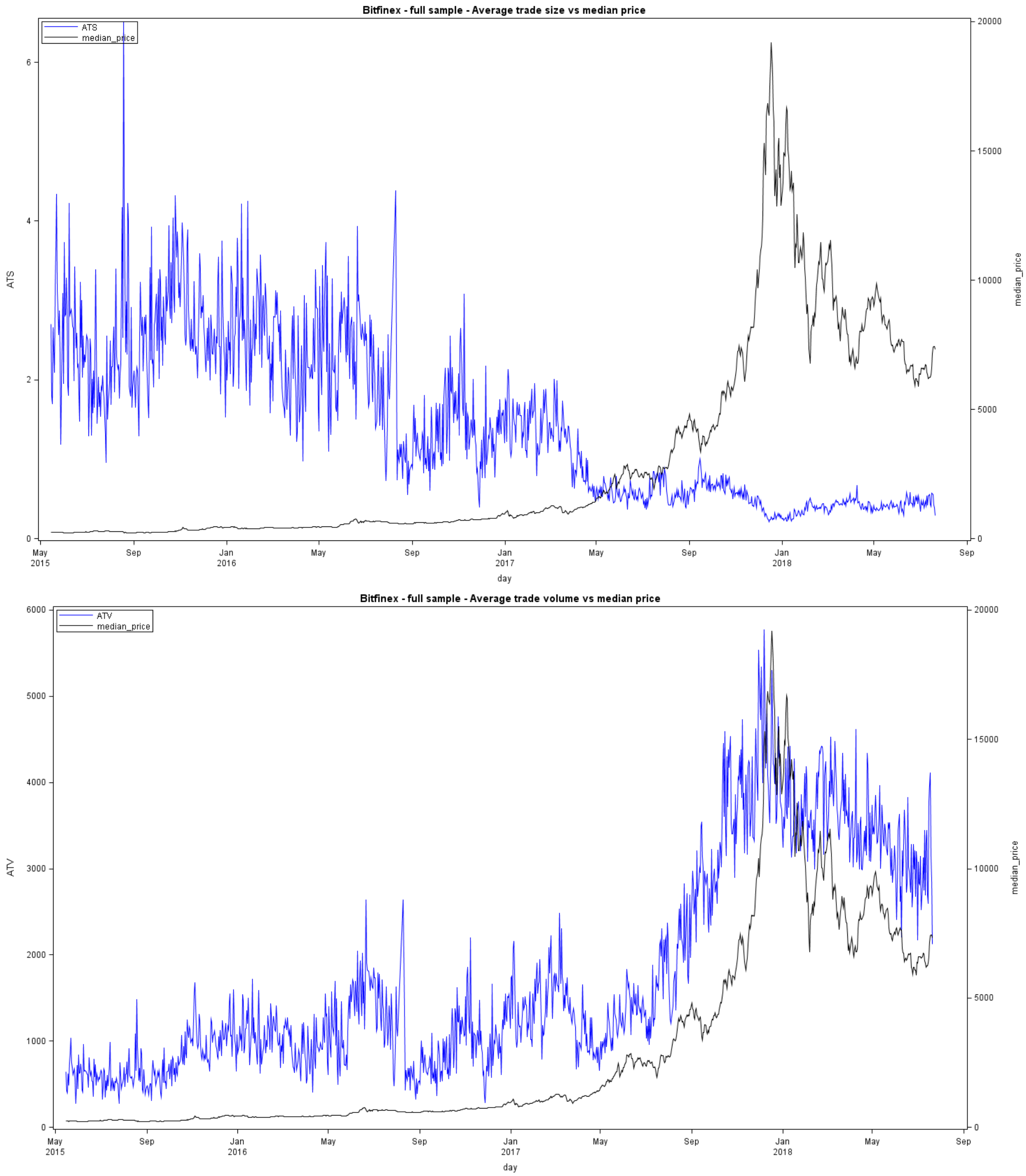
This figure represents the monthly market share of each platform for BTCUSD from May 2015 to July 2018.

Figure A2: BTCUSD - Platforms' monthly market share



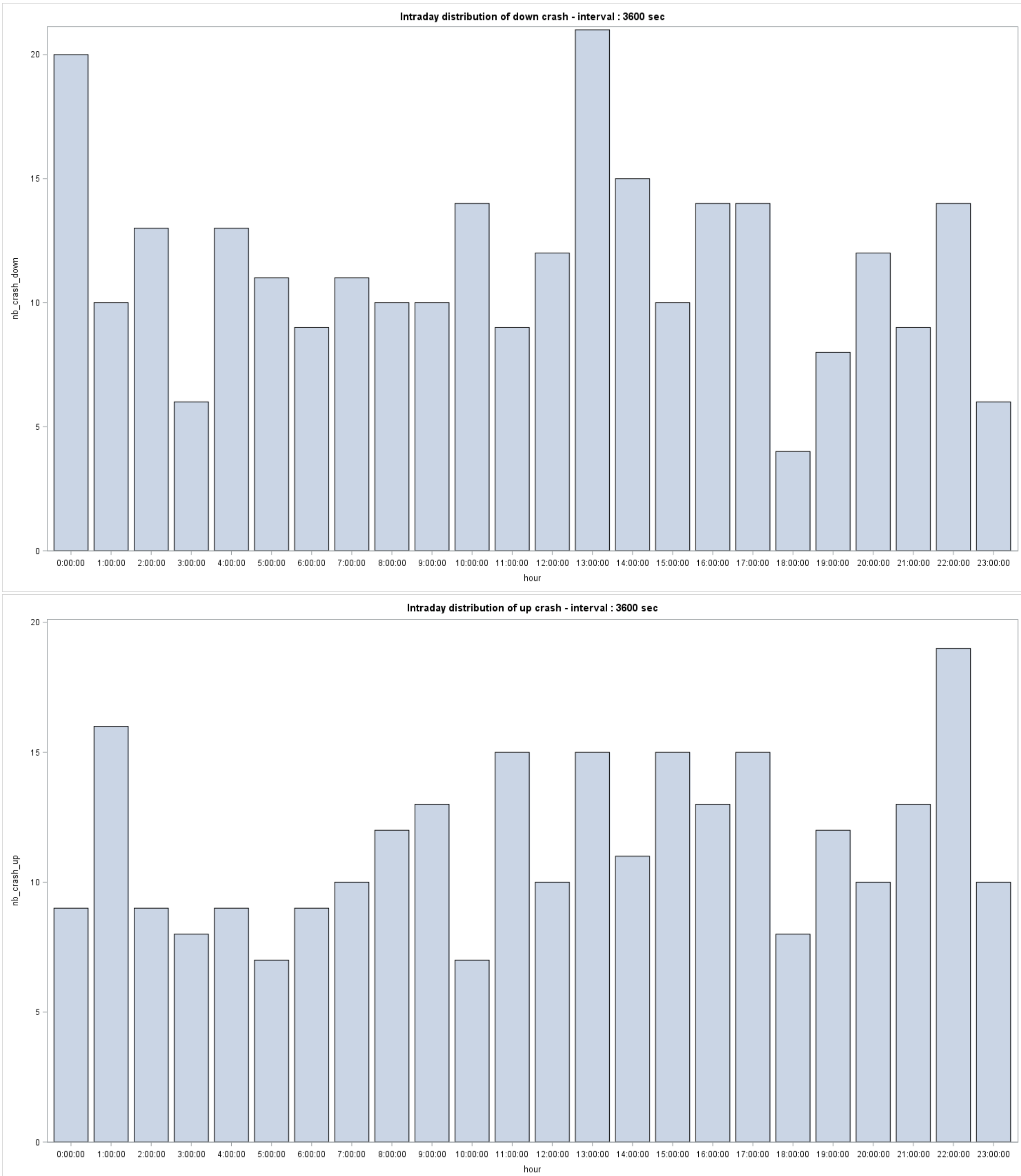
This figure represents the platforms' monthly market share for BTCUSD from July 2010 to September 2018.

Figure A3: ATS and ATV versus price



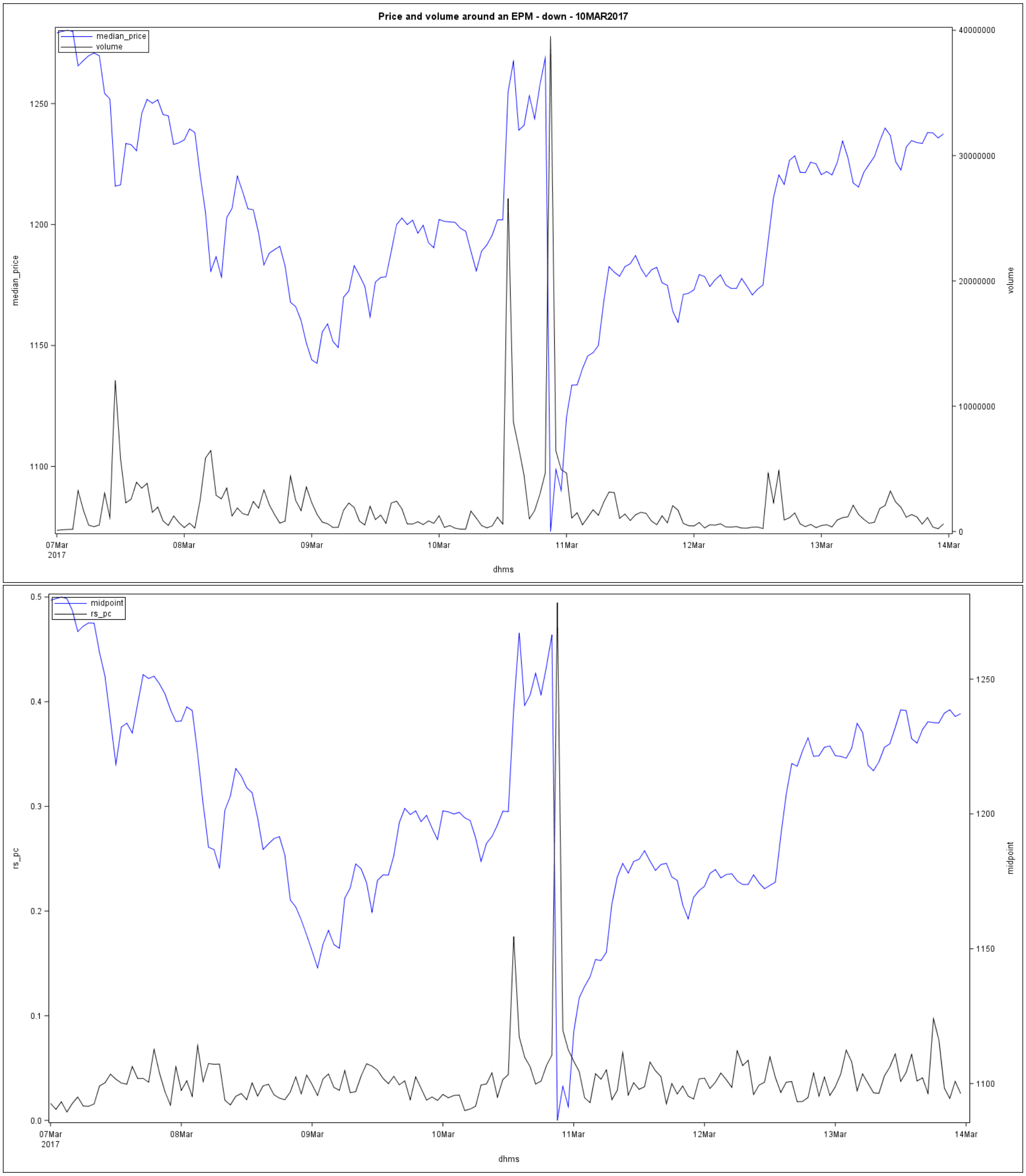
These figures represent the relationship between the daily median price (black line) and the average trade size (ATS, blue line) (above) and between the daily median price (black line) and the average trade volume (ATV, blue line) from May 2015 to July 2018.

Figure A4: Intraday distribution of EPMs



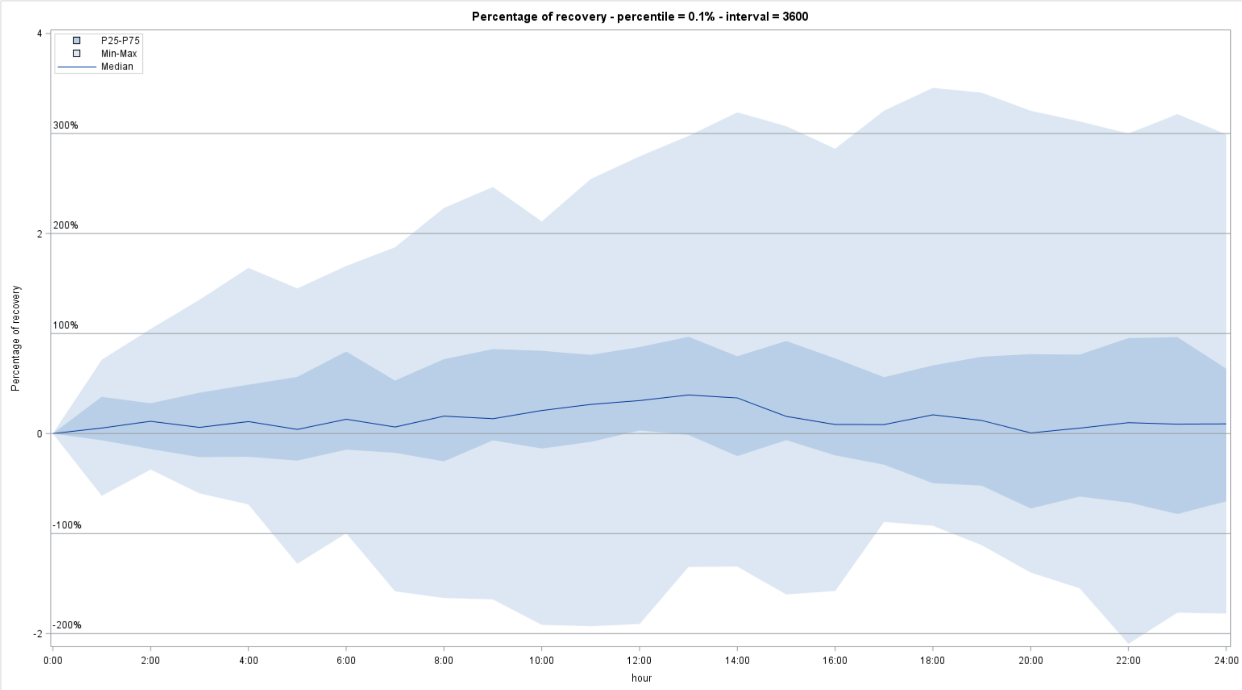
These figure represent the intraday distribution of EPM. We separate the graphs between down crashes (above) and up crashes (below).

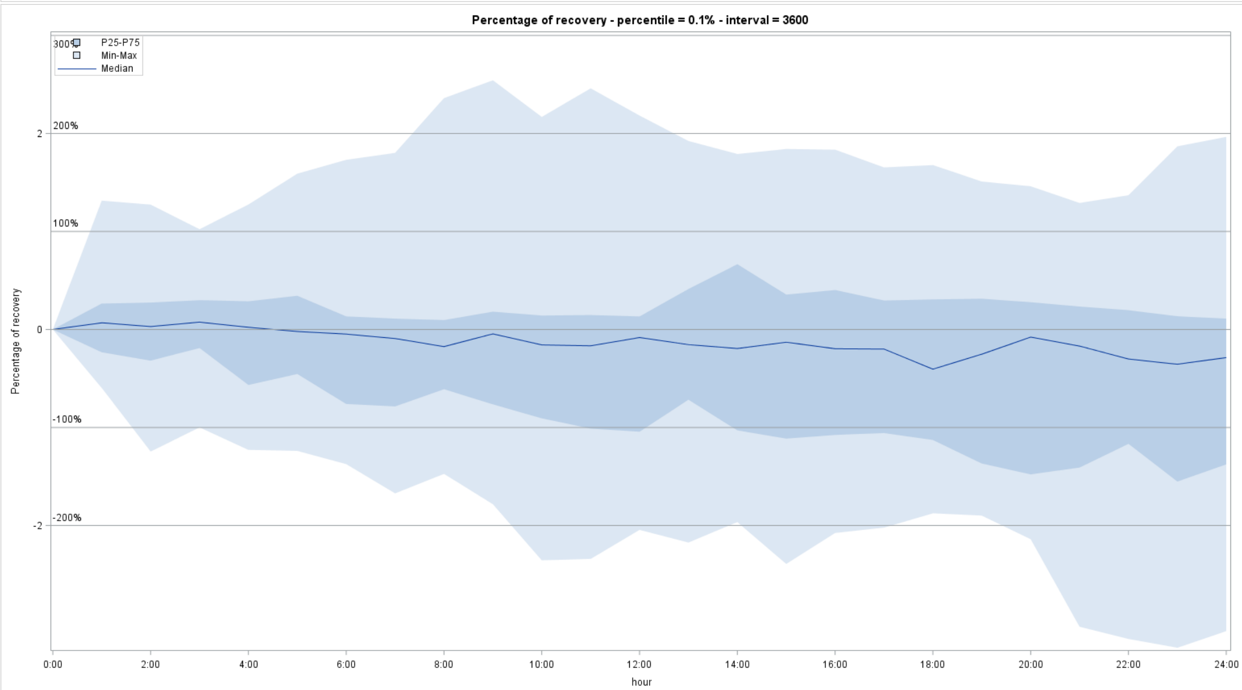
Figure A5: Illustrative example - relationship between price, trading volume, and liquidity



These figures represent the relationship between BTCUSD price and volume (above) and the relationship between BTCUSD price and relative spread (below) during an EPM.

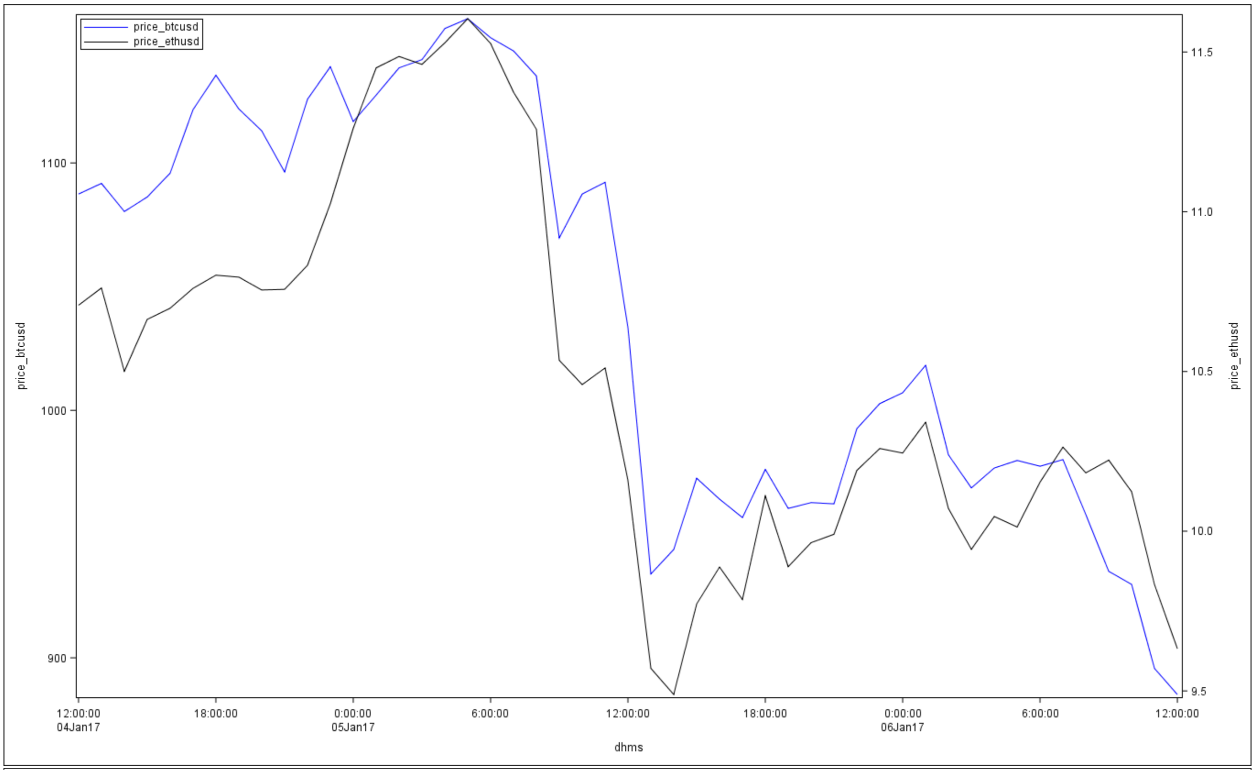
Figure A6: Percentage of recovery after an EPM

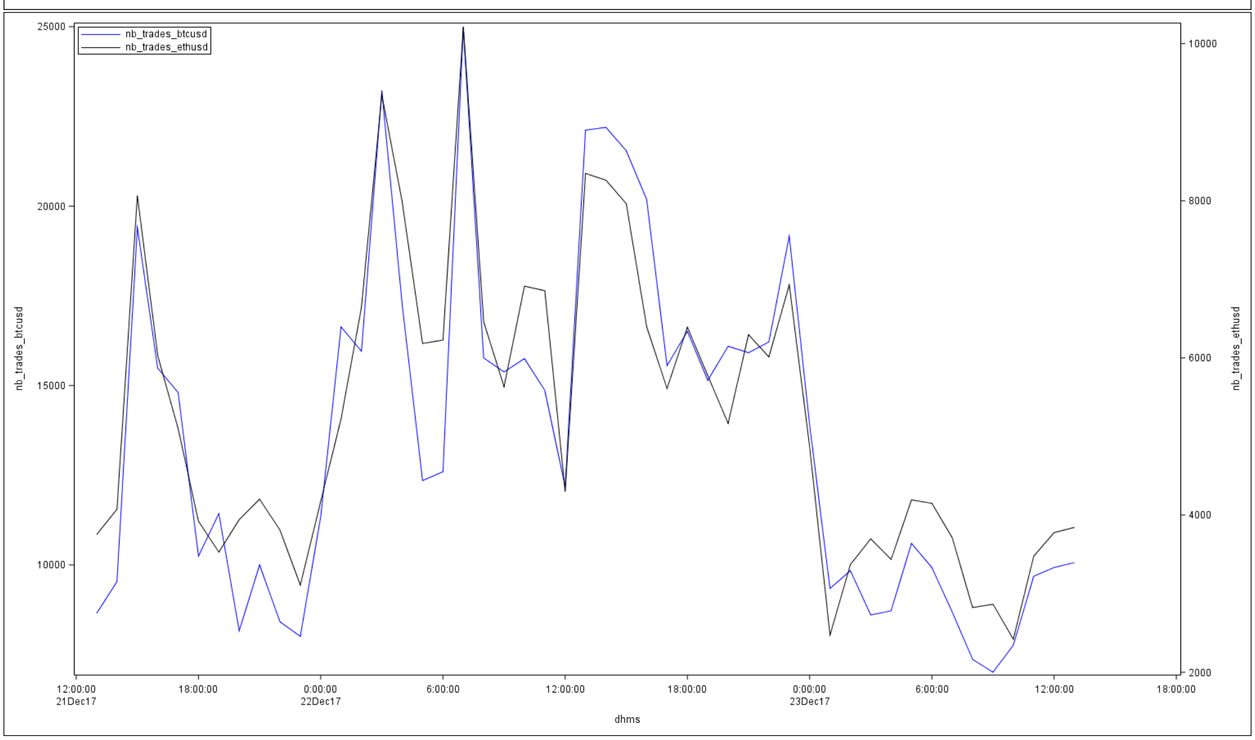




These figures represent the percentage of recovery conditionally on time, up to 24h after an EPM. We represent the median value, a confidence band between the 25^{th} and the 75^{th} percentile (...) and the confidence band between the minimum and the maximum recovery. We separate the graphs between down crashes (above) and up crashes (below).

Figure A7: BTCUSD vs. ETHUSD - price (above) and trading activity (below)





The above figure represents the evolution of BTCUSD (blue line) and ETHUSD (black line) prices during an EPM occuring in BTCUSD. The figure below represents the evolution of trading activity (in terms of number of trades), respectively for BTCUSD (blue line) and ETHUSD (black line).

B. Tables

Table A1: Definitions

Bitcoin
www.bitcoin.org: an innovative payment network and a new kind of money.
European Central Bank (2012, p. 13): a type of unregulated, digital money, which is issued and controlled by its developers, and used and accepted among the members of a specific virtual community.
Balcilar et al. (2017, p. 75): an open source software-based online payment system.
Hale et al. (2018): a cryptocurrency - a digital currency that is not backed by any tangible or intangible assets or intrinsic value.
Hayes (2017, p. 1309): the first and most popular of what has become known as cryptocurrencies, digital monetary and payment systems that exist online via decentralized, distributed networks that employ a shared ledger data technology known as blockchain coupled with secure encryption. (...) Bitcoin is an open source software-based online payment system that emerged in 2008-2009. Payments are recorded in a shared public ledger, known as the blockchain, using its own unit of account, which is also called bitcoin, symbolically represented as either BTC or XBT.
Baur et al. (2018, p. 177): a digital money within a decentralized peer-to-peer payment network. It is a hybrid between fiat currency and commodity currency without intrinsic value and independent of any government or monetary authority.
Brandvold et al. (2015): a decentralized peer-to-peer crypto-currency protocol.
Cryptocurrencies

Osterrieder and Lorenz (2017); Chu et al. (2017): a digital asset designed to work as a medium of exchange using cryptography to secure the transactions and to control the creation of additional units of the currencies.
Corbet et al. (2019): peer-to-peer electronic cash systems which allow online payments to be sent directly from one party to another without going through a financial institution.
Chu et al. (2017, p. 1): a digital asset designed to work as a medium of exchange using cryptography to secure the transactions and to control the creation of additional units of the currency.
Virtual currencies
European Central Bank (2015): a digital representation of value, not issued by a central bank, credit institution or e-money institution, which in some circumstances, can be used as an alternative money.
Digital currencies
Kristoufek (2013): A digital currency can be defined as an alternative currency which is exclusively electronic and thus has no physical form. It is also not issued by any specific central bank or government of a specific country and it is thus practically detached from the real economy.
Electronic currencies
Dwyer (2015, p. 91): an asset which can change hands from one person to another and is evidenced by a balance sheet that the owner of the currency keeps.
Deposits
Dwyer (2015, p. 91): money which is evidenced by an account at a bank and is a liability of that institution.

This table reports the most relevant definitions identified in the literature.

Table A2: Absolute Log returns - Bitfinex - BTCUSD

Platform	Crypto	P50	P75	P90	P95	P99	P99.9	P99.99
Panel A: BTCUSD								
bitfinex	btcusd	0.0030	0.0069	0.0138	0.0208	0.0431	0.1158	0.2597
bitflyer	btcusd	0.0023	0.0047	0.0087	0.0127	0.0244	0.0509	0.0633
bitstamp	btcusd	0.0032	0.0071	0.0143	0.0214	0.0443	0.1241	0.2774
bittrex	btcusd	0.0033	0.0067	0.0118	0.0168	0.0307	0.0604	0.0642
btcc	btcusd	0.0068	0.0189	0.0362	0.0538	0.0937	0.1462	0.1487
btce	btcusd	0.0033	0.0072	0.0151	0.0230	0.0505	0.1252	0.2595
cexio	btcusd	0.0038	0.0083	0.0157	0.0228	0.0408	0.0768	0.1030
coinbase	btcusd	0.0024	0.0057	0.0118	0.0177	0.0356	0.0873	0.2723
gatecoin	btcusd	0.0031	0.0082	0.0191	0.0320	0.0704	0.1335	0.1933
gemini	btcusd	0.0027	0.0066	0.0135	0.0199	0.0373	0.0749	0.1471
hitbtc	btcusd	0.0042	0.0092	0.0176	0.0243	0.0429	0.0756	0.1193
huobi	btcusd	0.0012	0.0055	0.0135	0.0203	0.0400	0.0909	0.1469
kraken	btcusd	0.0035	0.0083	0.0167	0.0243	0.0492	0.1164	0.2295
mtgox	btcusd	0.0050	0.0110	0.0243	0.0397	0.0970	0.2488	0.4104
okcoin	btcusd	0.0024	0.0056	0.0113	0.0167	0.0320	0.0733	0.1163
quoine	btcusd	0.0033	0.0074	0.0148	0.0217	0.0431	0.0947	0.1637
Panel B: .../USD								
bitfinex	bchusd	0.0073	0.0151	0.0279	0.0396	0.0751	0.1359	0.2250
bitfinex	btceur	0.0040	0.0094	0.0187	0.0264	0.0470	0.0816	0.1279
bitfinex	btcusd	0.0030	0.0069	0.0138	0.0208	0.0431	0.1158	0.2597
bitfinex	eosusd	0.0084	0.0173	0.0306	0.0428	0.0768	0.1297	0.1688
bitfinex	ethusd	0.0054	0.0110	0.0204	0.0287	0.0525	0.1018	0.1489
bitfinex	ltcusd	0.0052	0.0108	0.0211	0.0315	0.0699	0.1771	0.6194
bitfinex	xlmusd	0.0069	0.0130	0.0214	0.0275	0.0483	0.0845	0.1074
bitfinex	xmrusd	0.0071	0.0142	0.0258	0.0363	0.0632	0.1142	0.2136
bitfinex	xrpusd	0.0061	0.0131	0.0248	0.0365	0.0751	0.1527	0.2905
Panel C: BTC/...								
bitfinex	btceur	0.0040	0.0094	0.0187	0.0264	0.0470	0.0816	0.1279
bitfinex	btcusd	0.0030	0.0069	0.0138	0.0208	0.0431	0.1158	0.2597

This table reports different percentiles (i.e. 50 ; 75 ; 90 ; 95 ; 99 ; 99.9 ; 99.99) for absolute log-returns computed for an interval of 3600s for several combinations of platforms and cryptocurrencies. We compute returns based on the last midpoint of the interval. The returns are expressed in percentage.

Table A3: Descriptive statistics - Interval = 30 minutes

	Mean	Median	StDev.	N
Panel A: Full sample				
ATS	1.34	0.86	1.29	54,641
ATV	1,651.59	1,210.74	1,341.80	54,641
$Amihud$	0.00	0.00	0.00	54,272
$DEPTH$	35,092.52	14,068.03	58,799.94	49,309
$DEPTH5$	108,832.70	57,589.65	130,701.50	49,309
NT	800.92	243.00	1,343.49	55,817
OB_Imb	(2.23)	(0.59)	10.95	49,309
OB_Imb5	(0.00)	(0.00)	0.27	49,309
RS	0.04	0.02	0.18	49,309
QS	0.62	0.23	1.89	49,309
QT	562.26	268.27	907.39	55,817
T_Imb	(0.02)	(0.01)	0.34	55,817
VT	2,555,961.00	242,053.10	5,974,644.00	55,817
Panel B: P99				
ATS	1.24	0.52	1.49	544
ATV	3,621.66	3,742.53	1,475.09	544
$Amihud$	0.00	0.00	0.00	544
$DEPTH$	66,796.25	52,504.10	71,037.89	507
$DEPTH5$	232,448.60	208,047.00	178,732.50	507
NT	5,414.99	5,124.00	3,546.50	544
OB_Imb	(2.42)	(1.07)	4.49	507
OB_Imb5	0.03	0.03	0.17	507
RS	0.07	0.04	0.10	507
QS	3.46	2.47	3.62	507
QT	4,103.03	3,159.49	3,893.68	544
T_Imb	(0.05)	(0.07)	0.20	544
VT	23,118,856.00	20,782,696.00	19,450,263.00	544
Panel C: P99.9				
ATS	1.65	0.57	1.96	55
ATV	4,112.64	4,111.29	1,808.57	55
$Amihud$	0.00	0.00	0.00	55
$DEPTH$	93,349.59	77,098.35	97,795.27	52
$DEPTH5$	314,660.20	286,623.90	259,255.10	52
NT	7,794.13	8,072.00	5,001.37	55
OB_Imb	(3.00)	(1.37)	4.87	52
OB_Imb5	(0.02)	(0.04)	0.19	52
RS	0.15	0.07	0.25	52
QS	5.89	4.07	6.53	52
QT	7,811.32	5,298.33	8,009.07	55
T_Imb	(0.06)	(0.04)	0.19	55
VT	37,425,135.00	35,881,019.00	28,953,983.00	55

This Table reports descriptive statistics about our variables: average trade size (ATS), average trade volume (ATV), Amihud (2002)'s ratio ($Amihud$, depth at best quotes ($Depth$), depth at the 5 best quotes ($Depth5$), number of trades (NT), orderbook imbalance (OB_Imb), orderbook imbalance at the 5 best quotes (OB_Imb5), relative spread (RS), quoted spread (QS), quantities traded (QT), trade imbalance (T_Imb), and volume traded (VT). All variables are defined in section B. All variables are averaged over a 30-min interval. For each variable, we report the mean, median, standard deviation, and number of observations.

Table A4: Descriptive statistics - Interval = 15 minutes

	Mean	Median	StDev.	N
Panel A: Full sample				
ATS	1.31	0.77	1.45	108,295
ATV	1,620.52	1,175.41	1,415.35	108,295
$Amihud$	0.00	0.00	0.00	107,334
$DEPTH$	35,387.60	13,848.18	66,040.04	96,837
$DEPTH5$	109,645.20	57,576.98	144,625.20	96,837
NT	400.53	116.00	698.54	111,637
OB_Imb	(2.21)	(0.40)	13.85	96,837
OB_Imb5	(0.00)	(0.00)	0.30	96,837
RS	0.03	0.02	0.19	96,837
QS	0.62	0.22	2.90	96,837
QT	280.77	121.31	505.99	111,637
T_Imb	(0.02)	-	0.40	111,637
VT	1,279,778.00	109,213.00	3,180,836.00	111,637
Panel B: Extreme price movements (EPM_{99})				
ATS	1.28	0.54	1.57	1,074
ATV	3,674.23	3,751.35	1,612.75	1,074
$Amihud$	0.00	0.00	0.00	1,074
$DEPTH$	70,017.70	46,533.13	98,080.71	994
$DEPTH5$	247,179.30	205,731.90	225,085.10	994
NT	2,893.31	2,600.00	1,944.52	1,074
OB_Imb	(2.58)	(0.74)	7.81	994
OB_Imb5	0.03	0.02	0.21	994
RS	0.07	0.04	0.12	994
QS	3.59	2.40	3.80	994
QT	2,296.64	1,655.98	2,242.79	1,074
T_Imb	(0.06)	(0.06)	0.25	1,074
VT	12,583,467.00	10,230,958.00	11,226,586.00	1,074
Panel C: Extreme price movements ($EPM_{99.9}$)				
ATS	1.77	0.62	1.93	107
ATV	3,757.55	3,766.26	1,621.20	107
$Amihud$	0.00	0.00	0.00	107
$DEPTH$	56,729.03	39,823.48	54,896.18	104
$DEPTH5$	240,187.20	204,188.10	196,767.30	104
NT	3,920.16	3,723.00	2,405.88	107
OB_Imb	(5.44)	(1.07)	17.97	104
OB_Imb5	(0.01)	(0.02)	0.21	104
QT	4,379.61	3,041.51	4,239.04	107
RS	0.16	0.08	0.31	104
QS	5.43	3.53	6.18	104
T_Imb	(0.08)	(0.06)	0.24	107
VT	17,326,519.00	15,885,108.00	14,595,752.00	107

This Table reports descriptive statistics about our variables: average trade size (ATS), average trade volume (ATV), Amihud (2002)'s ratio ($Amihud$, depth at best quotes ($Depth$), depth at the 5 best quotes ($Depth5$), number of trades (NT), orderbook imbalance (OB_Imb), orderbook imbalance at the 5 best quotes (OB_Imb5), relative spread (RS), quoted spread (QS), quantities traded (QT), trade imbalance (T_Imb), and volume traded (VT). All variables are defined in section B. All variables are averaged over a 15-min interval. For each variable, we report the mean, median, standard deviation, and number of observations.

Table A5: LOGIT - BTCUSD - All exchanges

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.2429	***	0.1902	***
T_Imb_{t-1}	(0.0994)	***	(0.1157)	
R_{t-1}	0.3729	***	0.3158	***
RS_{t-1}	0.1275	***	0.0526	***
dummy_bitfinex	(4.9288)	***	(7.1830)	***
dummy_bitflyer	(4.9331)	***	(7.2400)	***
dummy_bitstamp	(4.9034)	***	(7.2480)	***
dummy_bittrex	(5.2129)	***	(18.4181)	
dummy_btcc	(6.0014)	***	(7.3472)	***
dummy_btce	(4.9107)	***	(7.2010)	***
dummy_cexio	(4.9910)	***	(7.1905)	***
dummy_coinbase	(4.9143)	***	(7.2343)	***
dummy_gatecoin	(5.4069)	***	(8.2867)	***
dummy_gemini	(4.9047)	***	(7.1563)	***
dummy_hitbtc	(4.9108)	***	(7.0596)	***
dummy_huobi	(4.4643)	***	(6.7594)	***
dummy_kraken	(4.9296)	***	(7.2533)	***
dummy_okcoin	(4.8542)	***	(7.1676)	***
dummy_quoine	(5.0271)	***	(7.4279)	***
N	222,009		222,009	
$N_{y=0}$	219,799	99.00%	221,788	99.90%
$N_{y=1}$	2,210	1.00%	221	0.10%
R^2	72.42%		74.65%	

This Table reports results of Equation 6. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), and platforms' fixed effects. All non-dummy variables are standardized at the platform-level. In Panel A, we estimate a LOGIT regression. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table A6: LOGIT - BTCUSD - All exchanges

Panel B: LOGIT with FIRTH correction

	$EPM99$		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.2427	***	0.1896	***
T_Imb_{t-1}	(0.0994)	***	(0.1159)	*
R_{t-1}	0.3728	***	0.3156	***
RS_{t-1}	0.1268	***	0.0531	***
dummy_bitfinex	(4.9262)	***	(7.1583)	***
dummy_bitflyer	(4.9097)	***	(7.0132)	***
dummy_bitstamp	(4.9008)	***	(7.2252)	***
dummy_bittrex	(5.1198)	***	(7.3979)	***
dummy_btcc	(5.8844)	***	(6.9381)	***
dummy_btce	(4.9078)	***	(7.1741)	***
dummy_cexio	(4.9797)	***	(7.0917)	***
dummy_coinbase	(4.9117)	***	(7.2086)	***
dummy_gatecoin	(5.3931)	***	(8.0525)	***
dummy_gemini	(4.9013)	***	(7.1298)	***
dummy_hitbtc	(4.9031)	***	(6.9948)	***
dummy_huobi	(4.4603)	***	(6.7203)	***
dummy_kraken	(4.9264)	***	(7.2273)	***
dummy_okcoin	(4.8512)	***	(7.1415)	***
dummy_quoine	(5.0225)	***	(7.3831)	***
N	222,009		222,009	
$N_{y=0}$	219,799	99.00%	221,788	99.90%
$N_{y=1}$	2,210	1.00%	221	0.10%
R^2	72.41%		74.64%	

This Table reports results of Equation 6. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), and platforms' fixed effects. All non-dummy variables are standardized at the platform-level. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table A7: LOGIT - Bitfinex - All cryptocurrencies

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.27	***	0.04	
T_Imb_{t-1}	(0.03)		(0.01)	
R_{t-1}	0.24	***	0.11	***
RS_{t-1}	0.04	***	(0.02)	
dummy_bchusd	(4.80)	***	(6.94)	***
dummy_btcusd	(4.80)	***	(6.89)	***
dummy_eosusd	(4.75)	***	(6.81)	***
dummy_ethusd	(4.82)	***	(7.07)	***
dummy_ltcusd	(4.68)	***	(7.22)	***
dummy_xlmusd	(4.78)	***	(7.50)	***
dummy_xmrusd	(4.77)	***	(7.06)	***
dummy_xrpusd	(4.77)	***	(6.82)	***
N	98,176		98,176	
$N_{y=0}$	97,160	98.97%	98,084	99.91%
$N_{y=1}$	1,016	1.03%	92	0.09%
R^2	70.09%		72.53%	
Panel B: LOGIT with FIRTH correction				
	EPM_{99}		$EPM_{99.9}$	
NT_{t-1}	0.27	***	0.04	
T_Imb_{t-1}	(0.03)		(0.01)	
R_{t-1}	0.24	***	0.11	***
RS_{t-1}	0.04	***	(0.02)	
dummy_bchusd	(4.79)	***	(6.88)	***
dummy_btcusd	(4.80)	***	(6.87)	***
dummy_eosusd	(4.74)	***	(6.76)	***
dummy_ethusd	(4.81)	***	(7.04)	***
dummy_ltcusd	(4.68)	***	(7.17)	***
dummy_xlmusd	(4.76)	***	(7.09)	***
dummy_xmrusd	(4.77)	***	(6.99)	***
dummy_xrpusd	(4.76)	***	(6.76)	***
N	98,176		98,176	
$N_{y=0}$	97,160	98.97%	98,084	99.91%
$N_{y=1}$	1,016	1.03%	92	0.09%
R^2	70.08%		72.52%	

This Table reports results of Equation 7. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), and cryptocurrencies' fixed effects. All non-dummy variables are standardized at the cryptocurrency-level. In Panel A, we estimate a LOGIT regression. In Panel B, we estimate a LOGIT regression with Firth (1993)'s correction. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.

Table A8: LOGIT - All platforms - All cryptocurrencies

Panel A: LOGIT				
	EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.	
NT_{t-1}	0.2723	***	0.2141	***
T_Imb_{t-1}	(0.0831)	***	(0.1047)	
R_{t-1}	0.3436	***	0.3093	***
RS_{t-1}	0.1609	***	0.0577	***
dummy_bitfinex	(4.7339)	***	(6.9349)	***
dummy_bitflyer	(4.7467)	***	(6.9816)	***
dummy_bitstamp	(4.7041)	***	(6.9980)	***
dummy_bittrex	(5.0442)	***	(18.1699)	
dummy_btcc	(5.8155)	***	(7.0837)	***
dummy_btce	(4.7146)	***	(6.9647)	***
dummy_cexio	(4.8184)	***	(6.9423)	***
dummy_coinbase	(4.7186)	***	(6.9863)	***
dummy_gatecoin	(5.2369)	***	(8.0445)	***
dummy_gemini	(4.7146)	***	(6.8977)	***
dummy_hitbtc	(4.7272)	***	(6.8018)	***
dummy_huobi	(4.2353)	***	(6.4709)	***
dummy_itbit	(4.7066)	***	(6.8919)	***
dummy_kraken	(4.7360)	***	(6.9998)	***
dummy_okcoin	(4.6539)	***	(6.9040)	***
dummy_quoine	(4.8394)	***	(7.2368)	***
dummy_bchusd	(0.2371)		(0.2814)	
dummy_btcusd	(0.1937)		(0.2734)	
dummy_eosusd	(0.1164)		(0.0410)	
dummy_ethusd	(0.1671)		(0.3304)	
dummy_ltcusd	(0.0496)		(0.7706)	
dummy_xlmusd	(0.1777)		(0.7966)	
dummy_xmrusd	(0.1745)		(0.3620)	
dummy_xrpusd	(0.1784)		(0.1801)	
dummy_btceur	(0.2699)		(0.4306)	
N	320,288		320,288	
$N_{y=0}$	317,055	98.99%	319,975	99.90%
$N_{y=1}$	3,233	1.01%	313	0.10%
R^2	72.38%		74.65%	

This Table reports results of Equation 7. The dependent variable is the occurrence of an EPM at time t and the independent variables include the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS), platforms- and cryptocurrencies' fixed effects. All non-dummy variables are standardized at the platform-cryptocurrency-level.

Table A9: LOGIT - interval = 1800s

	LOGIT				LOGIT (Firth)			
	EPM_{99}		$EPM_{99.9}$		EPM_{99}		$EPM_{99.9}$	
Variable	Coeff.		Coeff.		Coeff.		Coeff.	
α_0	-5.0051	***	-7.346	***	-5.0014	***	-7.3153	***
NT_{t-1}	0.5026	***	0.4323	***	0.5026	***	0.434	***
T_Imb_{t-1}	-0.0178		-0.0779		-0.0178		-0.0784	
R_{t-1}	0.1774	***	0.1732	***	0.1768	***	0.1702	***
RS_{t-1}	0.0264	*	0.0346		0.0318	***	0.0463	***
N	49,036		49,036		49,036		49,036	
$N_{y=0}$	48,528	98.96%	48,984	99.89%	48,528	98.96%	48,984	99.89%
$N_{y=1}$	508	1.04%	52	0.11%	508	1.04%	52	0.11%
R^2	1.81%		0.27%		1.82%		0.29%	

This Table reports results of Equation 5. The dependent variable is the occurrence of an EPM at time t and the independent variables include an intercept, the number of trades (NT), the trade imbalance (T_Imb), the absolute return (R), and the relative spread (RS). All variables, excepting the intercept, are lagged by one period and are standardized. We estimate with and without Firth (1993)'s correction. N is the number of observations, $N_{y=0}$ ($N_{y=1}$) is the number of non-events (events). We also report the R-square. *, **, *** indicate respectively statistical difference at the 10%, 5%, and 1% level.