

MIT Open Access Articles

Trading and arbitrage in cryptocurrency markets

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Makarov, Igor and Antoinette Schoar. "Trading and arbitrage in cryptocurrency markets." *Journal of Financial Economics* 135, 2 (February 2020): 293-319.

Published Version: <http://dx.doi.org/10.1016/j.jfineco.2019.07.001>

Publisher: Elsevier BV

Permanent Link: <https://hdl.handle.net/1721.1/130495>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: <http://creativecommons.org/licenses/by-nc-nd/4.0/>



Trading and Arbitrage in Cryptocurrency Markets

Igor Makarov^{a,*}, Antoinette Schoar^{b,*}

^a*London School of Economics*

^b*MIT Sloan, NBER, CEPR*

Abstract

Cryptocurrency markets exhibit periods of large, recurrent arbitrage opportunities across exchanges. These price deviations are much larger across than within countries, and smaller between cryptocurrencies, highlighting the importance of capital controls for the movement of arbitrage capital. Price deviations across countries co-move and open up in times of large bitcoin appreciation. Countries with higher bitcoin premia over the US bitcoin price see widening arbitrage deviations when bitcoin appreciates. Finally, we decompose signed volume on each exchange into a common and an idiosyncratic component. The common component explains 80% of bitcoin returns. The idiosyncratic components help explain arbitrage spreads between exchanges.

JEL classification: G12, G15, F38

Keywords: Cryptocurrencies, Bitcoin, Arbitrage, Price Impact, Capital Controls

*Igor Makarov: Houghton Street, London WC2A 2AE, UK. Email: i.makarov@lse.ac.uk. Antoinette Schoar: Corresponding Author, 62-638, 100 Main Street, Cambridge MA 02138, USA. Email: aschoar@mit.edu. We thank Yupeng Wang and Yuting Wang for outstanding research assistance. We thank seminar participants at the AEA, Brevan Howard Center at Imperial College, EPFL Lausanne, European Summer Symposium in Financial Markets 2018 Gerzensee, HSE Moscow, JOIM conference, LSE, Q Group Meeting, Nova Lisbon, and Yale as well as Jonathan Berk, Anastassia Fedyk, Adam Guren, Simon Gervais, Dong Lou, Peter Kondor, Gita Rao, Norman Schürhoff, Adrien Verdelhan and an anonymous referee for very helpful comments. Andreas Caravella, Robert Edström and Ambre Soubiran provided us with very useful information about the data. The data for this study was obtained from Kaiko Digital Assets.

1. Introduction

Cryptocurrencies have had a meteoric rise and a subsequent fall over the past few years. They are digital currencies that are built on blockchain technology that allows verification of payments and other transactions in the absence of a centralized custodian. Bitcoin, the most famous and earliest cryptocurrency, was originally introduced in a paper by Nakamoto (2008) and came into existence in 2009. Since then, the market for cryptocurrencies has evolved dramatically. Today more than 50 million active investors trade bitcoin and other cryptocurrencies on more than 100 exchanges worldwide. While significant attention has been paid to the dramatic ups and downs in the volume and price of cryptocurrencies, there has not been a systematic analysis of the trading and efficiency of cryptocurrencies markets. In this paper we attempt to fill this gap.

A number of features make the cryptocurrency market a unique laboratory for studying arbitrage and price formation. There are many nonintegrated exchanges that are independently owned and exist in parallel across countries. On an individual basis the majority of these exchanges function like traditional equity markets where traders submit buy and sell orders, and the exchange clears trades based on a centralized order book. However, in contrast to traditional, regulated equity markets, the cryptocurrency market lacks any provisions to ensure that investors receive the best price when executing trades.¹ The absence of such mechanisms increases the role of arbitrageurs who can trade across different markets, but any constraints to the flow of arbitrage capital can result in markets potentially being segmented. Looking across markets then helps us understand which frictions lead to market segmentation. It also allows to analyze regional differences across investors in their demand for cryptocurrencies and the correlation structure of price movements across regions.

In the following we show a number of new stylized facts about the price formation across cryptocurrency markets. We first present our findings and then provide a possible explanation for these facts. For this analysis we use tick data for 34 exchanges across 19 countries. First, we show that there are large and recurring deviations in bitcoin prices across exchanges that open up across different exchanges and often persist for several hours, and, in some instances, days and weeks.

Second, price deviations are much larger across countries (or regions) than within the same country. The large deviations exist even between countries with the most

¹For example, the US Securities and Exchange Commission (SEC)'s National Best Bid and Offer (NBBO) regulation in the United States requires brokers to execute customer trades at the best available prices across multiple exchanges.

liquid exchanges, such as the US, Japan, Korea, and, to a lesser extent, Europe. We calculate that the daily average price ratio between the US and Korea from December 2017 until the beginning of February 2018 was more than 15% and reached 40% for several days. This has been noted in the popular press as the “Kimchi premium”. Similarly, the average price difference between Japan and the US was around 10%, and between US and Europe about 3%. To provide a sense of the magnitude of the money left on the table, we calculate the daily profits that could have been achieved in this market. The daily amount of potential arbitrage profits was often more than \$75 million, and in the period from December 2017 to February 2018 we estimate a minimum of \$2 billion of potential total arbitrage profit.² In contrast, the price deviations between exchanges in the same country typically do not exceed 1%, on average.

Third, we find that deviations in bitcoin prices across countries are highly asymmetric. In countries outside the US and Europe, bitcoin typically trades at a premium relative to the US and almost never at a price below the US. In addition, there is significant co-movement in price deviations across countries: arbitrage spreads open up and close at the same time across countries.

Fourth, our analysis shows that price deviations occur during periods of a particularly quick appreciation of bitcoin prices. Since we show later that bitcoin prices react strongly to order flows, these periods also coincide with the times when there is a particularly strong increase in demand for bitcoin worldwide. To construct a measure of “buying pressure” in bitcoin markets, we take the difference between the actual log price of bitcoin in the US and its trend component, which we estimate using the Hodrick-Prescott filter. The bitcoin price in the US is a good proxy for the world market price of bitcoin. We then regress the deviations of a country’s bitcoin price relative to the US on our measure of “buying pressure”. This gives us a measure of the sensitivity of a country’s bitcoin price to changes in the world market price of bitcoin. We call this the bitcoin beta of a country. We show that the countries that, on average, have a higher premium over the US bitcoin price are also those with a higher bitcoin beta. So these countries respond more strongly in widening arbitrage deviations in times when buying pressure goes up in the US.

Our results thus show that the marginal investor outside the US and Europe is willing to pay more for bitcoin in response to positive news or sentiment. How can one explain this differences in valuation? We conjecture that they might reflect tighter

²Our approach allows us to abstract from any assumptions about price impact of additional arbitrage trades or the speed of convergence. Since we rely on trades that were executed on the exchanges, it also eliminates concerns about stale prices or illiquid exchanges. See Section 4.4 for details.

capital controls or weaker financial institutions in the countries outside the US and Europe. The marginal investor in a country with poorly functioning financial institutions or tighter capital controls might be willing to pay more for bitcoin since they would benefit more from the adoption of cryptocurrencies. Therefore, any news about the potential adoption of bitcoin technology (or any sentiment change) would increase the price in these countries more.

Of course, the price differences could only persist if capital markets are segmented or capital is slow to flow across borders. Controls on the flow of capital between regions reduce the efficient use of arbitrage capital. Arbitrage profits are realized by buying bitcoins in regions with low bitcoin prices, say the US, and selling in regions with high bitcoin prices, say Korea. This trade requires capital in the US and generates profits in fiat currency in Korea. If this profit cannot be repatriated seamlessly from Korea to the US, arbitrage capital can become “stuck” within a country and thus become scarce. It is important to note that capital controls are most relevant for the price of crypto to fiat currencies. In contrast, transactions between two cryptocurrencies should not be affected since they are explicitly designed to circumvent restriction to capital flows.

To test the importance of capital controls on fiat currencies, we analyze whether the positive correlation in arbitrage spreads between the countries we show above is explained by the level of openness of a country. If countries that are relatively closed have a higher convenience yield for bitcoin, then we should see their arbitrage spreads (relative to the world market price) move more closely together. Countries that are more open should not be correlated since any price deviation will be immediately arbitrated away. As measure of the pairwise tightness in capital controls between two countries, we take the product of the capital control index of the two countries constructed by [Fernandez et al. \(2015\)](#). This measure is zero if at least one of the countries is totally open and approaches one if both countries have very high levels of capital controls. In support of the idea that capital market segmentation is important in explaining arbitrages in bitcoin prices, we find that there is a significantly positive relation between the correlation of arbitrage spreads and capital controls. In other words, two countries that are both relatively closed to capital flows have a higher correlation in arbitrage spreads.

In further support of the idea that capital controls play an important role, we find that arbitrage spreads are an order of magnitude smaller between cryptocurrencies (say bitcoin to ethereum or to ripple) on the exact same exchanges where we see big and persistent arbitrage spreads relative to fiat currencies. For example, during the same period when the difference in the price of dollar to bitcoin between the US and Korea was more than 20%, the difference in the price of ethereum to bitcoin

was 3%, on average. Furthermore, the price of ethereum to fiat currencies shows as large an arbitrage spread across exchanges as the bitcoin market. We show similar patterns for the exchange rates between ripple and bitcoin or ethereum. Since the main difference between fiat and cryptocurrencies is the inability to enforce capital controls, our findings suggest that such controls contribute to the large arbitrage spreads we find across regions.³

Our findings suggest that there are significant barriers to arbitrage between regions and, to a lesser extent, even between exchanges in the same country. But the magnitude of the arbitrage spreads we show above are still surprising. While regulations in some countries make cross-border transfers in fiat currencies difficult for retail investors,⁴ industry reports suggest that large institutions should be able to avoid these constraints; see, for example, a recent International Monetary Fund (IMF) working paper by [Baba and Kokenyne \(2011\)](#).

We also conduct a number of robustness tests to show that mere transaction costs cannot explain the size of arbitrage spreads across exchanges since their magnitudes are small in comparison to the arbitrage spreads we show. Similarly, the governance risk of cryptocurrency exchanges being hacked or misappropriating client funds is also unlikely to explain these arbitrage spreads. Note that governance risk by itself does not predict the direction of any arbitrage spreads. Instead, one would expect that governance risk is correlated with lower trading volume and higher bid-ask spreads on an exchange since market participants should be hesitant to use such an exchange. To explain the large cross-border arbitrages we observe, the governance risk would have to disproportionately affect exchanges in one region versus another. This does not seem to be supported by the data. First, all regions we consider have liquid exchanges with similar volumes and bid-ask spreads. Second, we find large heterogeneity in the liquidity of exchanges within a region, but nevertheless arbitrage spreads are small between them. However, it is possible that ex ante concerns about the safety of bitcoin exchanges might make some arbitrageurs stay out of the market all together, which might ultimately explain why

³One additional factor to consider is that some companies manage exchanges in several countries. For example, Coinbase has operations in Australia, Canada, Europe, Great Britain, and the US. The order book of each of these exchanges are separate, and customers from different countries can usually only trade cryptocurrencies on their local exchange and in their local currency. However, the exchanges that operate across regions might be able to arbitrage and potentially circumvent some capital controls. While we do not find a consistent impact of having overlapping exchanges on the arbitrage spreads of countries, it is possible that the existence of this channel adds noise to the estimation of capital controls.

⁴For example, in Korea, local residents and companies moving more than \$50,000 out of the country in a single year must submit documents to authorities proving their reasons for the transfers, which may not always be approved.

arbitrage capital is limited. Of course, increasing professionalization, better governance of the cryptocurrency market, and innovations over time might reduce these constraints to arbitrage. For example, moving trading completely into crypto space by substituting fiat currencies with their digital counterparts, such as tether or Circle's digital version of the US dollar, can diminish the role of capital controls.

Finally, we study the price impact in the bitcoin market to provide an estimate of how much capital is required to close the arbitrage spreads we show above. We focus on Kyle (1985)'s lambda that measures the price pressure of net order flow. Previous research in other asset classes attributes the price pressure of net order flow to price discovery, but in the cryptocurrency market it is less obvious whether there are traders who are more informed than others and what the nature of the information is. Nevertheless, we show that a strong positive relation also exists between net order flows and prices in the cryptocurrency market. A common way to estimate the impact of net order flow is to regress returns over a particular time period on the signed volume of trades during the same period. The complication in the bitcoin market is that the same asset is traded simultaneously on multiple exchanges. When forming their demand, investors might not only look at prices on their own exchange but also take into account prices on the other exchanges where bitcoin is traded. Therefore, we decompose signed volume and returns on each exchange into a common component and an idiosyncratic, exchange-specific component. We use factor analysis to extract the common factors from data at five-minute, hourly, and daily frequencies. The common component of signed volume explains about 50% of the variation in returns at the five-minute and hour level and up 85% at daily level. The price pressure at the daily level is mostly permanent: buying 10,000 bitcoins raises returns by about 4%.

To investigate the role of signed volume in explaining price deviations across exchanges, we show that exchange-specific residuals of signed volume are significant at explaining variation in exchange-specific residuals of returns at the five-minute and hour level. We also show that when the price on any exchange deviates above (below) from the average price on other exchanges, subsequent returns on this exchange are predicted to be lower (higher) than the returns on other exchanges. These results show that arbitrage spreads open up in periods when there are differential price pressures through idiosyncratic signed volume on one exchange relative to another. The price deviations are not arbitrated away immediately, but they do predict subsequent relative returns on exchanges.

Our paper is related to several streams of the literature. Research on cryptocurrencies in finance and economics is still in its beginning. The majority of papers in this literature focuses on the potential real effects of cryptocurrencies as a payment

and transaction mechanism. [Ciaian, Rajcaniova, and Kancs \(2016\)](#), [Harvey \(2016\)](#), [Bohme et al. \(2015\)](#), and [Raskin and Yermack \(2017\)](#) provide a broad perspective on the economics of cryptocurrencies and the blockchain technology they are built upon. [Athey et al. \(2016\)](#) and [Pagnotta and Buraschi \(2018\)](#) propose models of the valuation of digital currencies. [Cong, He, and Li \(2019\)](#), [Easley, O'Hara, and Basu \(2017\)](#), and [Huberman, Leshno and Moallemi \(2017\)](#) study Bitcoin mining fees and the incentives of miners in equilibrium. We view our paper as complementary to this literature. To our knowledge, we are the first to provide a systematic empirical study of trading and price formation in cryptocurrency markets using transaction-level data.

Our paper is also linked to the limits of arbitrage, which argues that prices can deviate from law of one price even in the presence of arbitrageurs; see, e.g., [DeLong et al. \(1990\)](#), [Gromb and Vayanos \(2002\)](#), and [Gromb and Vayanos \(2018\)](#). On the empirical side, our paper is closest to the studies that analyze deviations from one price in different markets. In particular, [Rosenthal and Young \(1990\)](#) and [Froot and Dabora \(1999\)](#) study “Siamese twin” companies. They show that prices of two types of shares, which are traded in different markets but have identical claims on the cash flows and assets of the same company, can nevertheless substantially deviate from each other. Similar to [Rosenthal and Young \(1990\)](#) and [Froot and Dabora \(1999\)](#), we show that bitcoin and other cryptocurrencies can be traded at vastly different prices on different exchanges. The deviation from the law of one price is even more striking in the case of cryptocurrencies, since unlike shares that are traded within specific countries, bitcoins can be transferred to any market. As a result, typical explanations such as tax-induced investor heterogeneity or index membership do not apply in this case. On a broader level, our paper is also linked to the market segmentation literature; see, e.g., [Bekaert et al. \(2011\)](#). Similar to this literature, our results suggest that capital controls and the development of financial markets can be important at explaining the differences in the marginal valuation of investors across countries.

Finally, our paper is also related to research that shows a strong positive relation between asset prices and net order flow in “traditional” financial markets. For example, [Evans and Lyons \(2002\)](#), [Berger et al. \(2008\)](#), and [Fourel et al. \(2015\)](#) look at foreign exchange markets; [Brandt and Kavajecz \(2004\)](#) at US Treasury markets; [Deuskar and Johnson \(2011\)](#) at the S&P 500 futures market; and [Chordia, Roll, and Subrahmanyam \(2002\)](#), [Goyenko, Holden, and Trzcinka \(2009\)](#), and [Hendershott and Menkveld \(2014\)](#) for NYSE stocks. These papers suggest that order flow imbalances typically explain about 15%-30% of the day-to-day variation of stock returns or treasury yields and up to 50% of foreign exchange returns. We show that a very strong positive relation exists in cryptocurrency markets as well. But the R -squared that we show for cryptocurrency

markets are significantly higher, up to 85%.

To estimate price impact, we follow most closely the econometric approach of [Hasbrouck \(1995\)](#) and [Hasbrouck and Seppi \(2001\)](#), who study common factors in stock returns and order flows in a cross-section of US stocks. Similar to [Hasbrouck \(1995\)](#), we rely on the idea that prices across different markets are cointegrated and thus cannot diverge too far from each other. This allows us to decompose the price on each exchange into the common component (what Hasbrouck calls the “implicit efficient price”) and an exchange-specific deviation from the common component. One notable difference of our approach from these earlier papers is that we are interested in estimating the magnitude of the price impact of signed volume on the common component and an exchange-specific component. For this purpose we combine factor analysis with the price decomposition in [Hasbrouck \(1995\)](#) and impose appropriate constraints on factor loadings and weights.

The rest of the paper is structured as follows. Sections 2 and 3 explain the data and provide summary statistics of volume and returns. Section 4 presents the results of the arbitrage index and arbitrage profits within and across regions. In Section 5 we discuss the correlations structure of price deviations and in Section 6 we show arbitrage spreads across other crypto-currencies. In Section 7 we discuss where arbitrage spreads might arise and estimate the model of price pressure. Finally, Section 8 discusses the implication of our findings for arbitrage dynamics in the cryptocurrency market, and Section 9 concludes.

2. Data description

The main data for this project are tick-level trading data obtained from Kaiko, a private firm that has been collecting trading information about cryptocurrencies since 2014. The Kaiko data cover the 17 largest and most liquid exchanges: Binance, Bitfinex, bitFlyer, Bithumb, Bitstamp, Bitbox, Bittrex, BTCC, BTC-e, Coinbase, Gemini, Huobi, Kraken, OkCoin, Poloniex, Quoine, and Zaif. Besides Bitcoin, which is the first and most famous cryptocurrency, Kaiko also provides trading information on other coins such as ethereum, ripple, and tether. We restrict our attention to the three most liquid and largest cryptocurrency markets: bitcoin (BTC), ethereum (ETH), and ripple (XRP). We focus our analysis on the period from to January 1, 2017, to February 28, 2018. This choice is motivated by the market liquidity. Prior to these dates, the liquidity in crypto markets was significantly lower than in later periods.

Kaiko obtains the data by querying APIs provided by the exchanges. The variables contained in the data are the time stamp of the transaction at the second and mil-

lisecond levels, the price at which the trade happened, the amount of the trade, and an indicator whether the trade was buy or sell initiated. To get information about the bid-ask spread, we also use Kaiko's order book data, which are obtained by querying the Application Programming Interface (API) of the exchanges and taking snapshots of their order books at the minute frequency.

In addition we use data from Bitcoincharts.com, a public website that also provides tick-level data by querying the API of exchanges at a second level. The data from Bitcoincharts.com compliment the universe of exchanges and geographic regions covered by Kaiko. Unlike data from Kaiko, data from Bitcoincharts.com have only the trade price and the amount of trade but not the direction of trade and are restricted to bitcoin transactions. Therefore, when exchange data are available from both Kaiko and Bitcoincharts.com, we use data from the former.⁵ Finally, the data from Bitso, Koinim, and Mercado are obtained by contacting the exchanges directly.

In total, we have data from 34 exchanges across 19 countries. In the majority of cases, the country of exchange operation defines the main fiat currency used as a base currency. For example, the main trading platform of bitFlyer, a Japanese exchange, offers trading of BTC to Japanese yen. However, a number of larger exchanges, Bitstamp, Bitx (Luno), Coinbase, Kraken, and Quoine, have operations across different countries and regions. The order book of each of these exchanges is separate, and customers from different countries can usually only trade cryptocurrencies on their local exchange and in their local currency. For example, Coinbase has operations in Australia, Canada, Europe, Great Britain, and the US. Coinbase allows US citizens and residents to open up a bitcoin trading account in the US and trade bitcoin against USD. But an investor in Australia would only be able to open an account and trade bitcoin against Australian dollars. The same rule applies to investors of Luno. Quoine also allows trading only in one base currency, which is, by default, the home currency of the investor's country of residence selected at registration. The base currency can be changed every six months. Exceptions are Bitstamp and Kraken, which allow trading against multiple currencies independent of the location of the customer. While this discussion suggests that in the majority of exchanges customers cannot easily arbitrage across fiat currencies, the exchanges that operate across regions might be able to arbitrage across regions and potentially circumvent some capital controls. Our analysis below suggests that overlapping exchanges do not have a consistent impact on the correlation of arbitrage spreads across countries. But it is possible that the existence

⁵The list of exchanges available both in Kaiko and Bitcoincharts.com include Coinbase, Bitstamp, bitFlyer, Bitfinex, Kraken, OkCoin, and Zaif. We compare the data across the two sources and find that the differences are small and infrequent.

of this channel adds noise to the estimation of capital controls.⁶

Bitstamp and Kraken are two exchanges that allow trading against USD, British pounds, or euros independent of the customer's location, as stated on their websites. However, when we called the exchanges to ask how it works, they confirmed that citizens of one country can only open accounts in their local currency. In addition, for all the exchanges that operate in multiple locations, their order books of bitcoin against each of the fiat currencies are completely separate.

The data on the exchanges used in the paper and markets they operate in are presented in Table 1. The choice of particular exchanges for each country is dictated by liquidity. For example, Coinbase has the largest trading volume in USD. The trading volumes in euro and GBP are also quite liquid, but the volume in Australian dollar is very small and prices are often stale. Therefore, we do not include data for the price of bitcoin to AUD on Coinbase in our analysis.

For the main analysis of the paper, where we establish the size of price deviations across markets and estimate price impact of trades, we focus only on the most liquid exchanges covered by Kaiko.⁷ In the second part, which looks more deeply into the role of capital market segmentation for the dynamics of arbitrage spreads, we use the extended data set across all 34 exchanges.

We classify the exchanges in Kaiko by region as follows:

- a. China: OkCoin, Huobi, and BTCC (base currency: Chinese yuan)
- b. Japan: bitFlyer, Zaif, and Quoine (base currency: Japanese yen)
- c. Korean: Bithumb (base currency: Korean won)

There are four major exchanges available in the US and Europe: Coinbase, Kraken, Bitstamp, and Gemini. The three largest of them, Coinbase, Bitstamp, and Kraken, allow trading of BTC both in US dollars and euros. But access to US citizens typically is in US dollars, while European citizens would open an account in euros. We therefore classify the euro trades on Bitstamp, Coinbase, and Kraken as belonging to Europe and the US dollar trades on Coinbase, Kraken, Bitstamp, and Gemini as belonging to the US.

- e. US: Coinbase, Kraken, Bitstamp, and Gemini (base currency: US dollar)
- f. Europe: Kraken, Coinbase, and Bitstamp (base currency: euro)

⁶See Table A1 in the appendix for a detailed list of exchanges that operate in several regions.

⁷We exclude BTC-e from this analysis because of the attempted closure of BTC-e by the US Justice Department during our sample period. We also exclude Bitbox because it has lower liquidity than the other Japanese exchanges: bitFlyer, Zaif, and Quoine.

The remaining three exchanges, Poloniex, Binance, and Bittrex, only allow trading between different cryptocurrencies and not fiat currencies. The main base currency used on these exchanges is tether. We also include Bitfinex in this set since it started using tether as its base currency in March 2017. Tether (USDT) is a cryptocurrency that was created in 2014; each token is supposed to be backed by one US dollar. The aim was to create a cryptocurrency that facilitates the digital transfer of fiat currencies with the stability of the US dollar. Tether has been highly traded since being used as a base currency for the above exchanges, with a value very close to the dollar.⁸

Daily exchange rate data are obtained from Bloomberg for the exchange rate pairs between Japanese yen and USD, Korean won and USD, and euro and USD. We also obtain hourly exchange rates for Euro: USD; these are reported as the first minute of a given hour.⁹

We found a number of data coding errors while cleaning up the data. These adjustments will be helpful for any researcher using the data in the future. First, we found that the time stamp on Korean exchanges, Bithumb, is reported in local Korean time and not in Coordinated Universal Time (UTC). The rest of the exchanges follow the UTC time convention. Second, we found that Bithumb and Quoine also seem to have an error in the convention of signing volume: sell-originated volume seems to be reported as buy-originated volume and vice versa. We confirm that this is the case by looking at the price impact (λ). For all exchanges, except for Bithumb and Quoine, λ is positive, while for Bithumb and Quoine the estimated λ is negative, irrespective of the time period of estimation. When we flip the sign, the λ , of course, becomes positive and looks similar to the coefficients estimated for the other bitcoin exchanges (see Section 7).

We also extensively clean the data for outliers and stale prices on days where an exchange was closed or experienced major system problems. The code can be obtained from the authors.

⁸However, since December 2017 there have been concerns about the transparency of Tether Limited, the company that is backing the 1:1 dollar claim of tether. The US Commodity Futures Trading Commission sent subpoenas to Tether Limited on December 6, 2017.

⁹We thank Adrien Verdelhan for sharing these data with us.

3. Summary statistics

3.1. Volume

We first show the total volume of bitcoin trading to their base currencies across the 15 exchanges in Kaiko. The choice of the base currency depends on the geographical focus of an exchange and can be found in Section 2. For example, for a Japanese exchange we would use the trading volume of bitcoin to Japanese yen. Several of these exchanges also allow trading of bitcoin to other cryptocurrencies such as ethereum. We do not include this volume here since the liquidity of these exchange rate pairs is often very limited. We analyze the trading in bitcoin to other coins in more detail below. Fig. 1, Panel A shows the average daily trading volume, averaged over the week, across all exchanges from January 2017 to the end of February 2018. We see that the daily trading volume was about 200,000 bitcoins across all exchanges at the beginning of 2017, which increases to almost 400,000 bitcoins in December and January 2018. We also see that a significant amount of volume is in tether, Japan, Korea, and the US.

[Fig. 1 About Here]

In Panel B of Fig. 1 we see, however, that the volume was an order of magnitude higher during 2016 and was driven by Chinese exchanges. Trading on bitcoin exchanges in the rest of the world combined constituted only a small amount in comparison. The high volume on Chinese exchanges in 2016 was reportedly because these exchanges had no trading fees and allowed margin trading with very high leverage. However, starting from January 2017, the People's Bank of China began exerting significant pressure on local exchanges to curb speculation and price volatility. In response, the exchanges implemented fixed trading fees of 0.2% per trade during that period and started migrating operations overseas or using peer-to-peer platforms.

The US and Japanese exchanges each have about 20% by the end of February 2018, European and Korean have around 10%. The graph also shows that the four fastest growing bitcoin exchanges are the tether-based exchanges: Bitfinex, Binance, Poloniex, and Bittrex. Many industry observers believe that some of the volume from China moved to trading on these exchanges.¹⁰

¹⁰A number of exchanges have been suspected of artificially inflating their volume or engaging in wash trading, especially Chinese exchanges. For this reason, we do not include these exchanges in our analysis. In addition, as we will show below, there is a strong common component in price impact and highly correlated price movements across exchanges even at a high frequency. This reduces the concern that the prices reported on the exchanges in our data are driven by noise.

In Table 2 we provide summary statistics on trading volume, number of trades, average trade size, and bid-ask spreads for each of the 15 exchanges in our data. We break out the statistics for the two subperiods, January 1 to July 31, 2017, and then August 1, 2017, to February 28, 2018, to reflect the significant changes in the bitcoin market. In the first half of the year, average daily trading volume of BTC in US dollars ranges from \$2.5 million to about \$20 million. But starting from August 2017 to the end of February 2018, the average daily trading volume increases significantly. On the largest exchanges, such as Bitfinex, Binance, or Coinbase, the average daily trading volume is around \$445 million, \$224 million, and \$180 million, respectively. In comparison the smaller exchanges typically have a daily volume of around \$45 million. This is more than a tenfold increase in trading volume for most exchanges. The only exceptions are the three Chinese exchanges, Okcoin, BTCC, and especially Huobi, which saw a dramatic drop in trading volume after January 2017.

The data for Binance and Bithumb are only available for the second half of 2017. While the daily trading volume increased significantly over this time period, it is still small relative to major forex markets. For example, the average daily volume of USD to euro is \$575 billion. But the magnitude is comparable to trading in a large firm like Netflix, which has an average daily volume of about \$3 billion.

A similar picture emerges when we look at the number of trades across exchanges, which also increases significantly. On most exchanges the number of trades almost triples between the first half of 2017 and the period from August 2017 to February 2018. Some of the biggest increases are on US and Japanese exchanges. Coinbase, the largest US exchange, went from 28,000 trades to about 84,000 trades a day. Similarly, bitFlyer, the largest Japanese exchange, increases from 32,000 trades per day to about 84,000. In comparison some of the smaller exchanges, such as Korbit or Gemini, on average have only 16,000 or 23,000 daily trades, respectively. Finally, the bid-ask spread on most of the exchanges is remarkably tight. The bid-ask spread, on average, is about ten basis points, and on the most liquid exchanges it is about two to three basis points.

3.2. Prices and returns

We now show the price dynamics of Bitcoin. Our data confirm the steep increase in the bitcoin price from January 2017 to January 2018, which has drawn a lot of attention in the popular press. The price rose from less than \$1,000 to almost \$20,000 at the end of 2017, with an especially rapid acceleration in the price of bitcoin after November 2017. The price fell back to just below \$10,000 by the end of February 2018.

Thus, from January 1, 2016, to February 28, 2018, the return on bitcoin is about 900%.

Table 3 shows the higher moments of bitcoin returns at the daily, hourly, and five-minute level from January 1, 2017, to February 28, 2018. These statistics are calculated by averaging the corresponding moments across all available exchanges. For each frequency we report the annualized standard deviation, skewness, and kurtosis of returns, as well as the autocorrelation and cross-correlation across exchanges.

Column (1) of Table 3 reports the standard deviation. We see that the volatility of returns is very high. Even at the daily frequency, the annualized standard deviation is 107%. In comparison, the annualized standard deviation of Nasdaq, from 1985 to 2017, is 18%. However, the kurtosis at the daily frequency is 3.86, which is not too far from that of the normal distribution. The daily returns are positively skewed, which is perhaps not very surprising given the steep increase in the price of bitcoin over the considered time period. Columns (4) through (6) show the autocorrelation in returns for one, two, and three lags. We can see that even at the five-minute frequency the autocorrelations are small, which shows that there is little predictability in the market.

Finally, in column (7) we report the average cross-correlation of returns. We average across all the pairwise correlations but take out the diagonal (i.e., the autocorrelation of an exchange with itself). We see that at the five-minute level the correlation between exchanges is quite low, only 57%, while at higher frequencies the correlation increases: it is 83% at the hourly level and 95% for daily returns. These results are similar to what is observed in other well-established markets; see, for example, [Budish, Cramton, and Shim \(2015\)](#). However, if in equity markets the break of correlations happens at millisecond frequencies, then here it is already present at the minute levels.

The lower correlations at higher frequencies point to the existence of price deviations between exchanges. In the following section we investigate the existence of arbitrage opportunities in more detail.

4. Arbitrage

4.1. Arbitrage index

The low cross-correlation in returns across exchanges, which we computed in the previous section, already suggests that the crypto market is far from being efficient. To show the amount of price dispersion between exchanges at a given point in time, we form an arbitrage index that compares the maximum difference in prices between exchanges. We start by calculating this arbitrage index at the minute level. For this purpose, for a given minute we first compute the volume-weighted average price in that

minute at each exchange,¹¹ and then we take the maximum price across all exchanges and divide it by the minimum price. Finally, we average the arbitrage index at the daily level to reduce the impact of intra-day volatility.

If the markets were completely integrated and arbitrage free, then the arbitrage index should be very close to one at all times. We first report the arbitrage index across all but the Chinese major exchanges. We exclude Chinese exchanges since the significant government interventions and trading restrictions on these exchanges could potentially make their prices less liquid and integrated with those on other exchanges. Fig. 2 shows that during the period from January 1, 2017, to February 28, 2018, there is significant variability during the year. Remarkably there are several months during the year where the index stays at about 1.5, for example, May and June 2017, as well as December 2017 to mid-February 2018.

[Fig. 2 About Here]

4.2. Arbitrage index within geographical regions

To understand where the price deviations occur, we decompose the arbitrage index into price differences within regions versus across regions. We start by looking separately at the arbitrage index within each of the major regions, where bitcoin trading is prevalent and we have more than one exchange (i.e., Europe, Korea, Japan, and the US). For each region, we have at least three exchanges; the only exception is Korea, where we only have two exchanges, Bithumb and Korbit.

In Fig. 3, Panel A we report the arbitrage index for the four major exchanges operated in the US (Bitstamp, Coinbase, Gemini, and Kraken) from January 2017 to February 2018.¹² The calculations follow the exact same approach as above. We see that for most of the year the arbitrage index is very low; the average price dispersion is below 2%. But there are a few weeks in March, June, and December where the arbitrage index is around 1.04. While the arbitrage spreads within the region are small compared to the total arbitrage index reported above, they are still large in comparison to more traditional financial markets.

We then repeat the same analysis for the arbitrage index among European exchanges for the period from January 2017 to February 2018. Again, we see in Fig. 3

¹¹To compute the volume-weighted average price in a given period, we multiply each transaction price in the period by the respective transaction amount, sum them, and divide by the total volume of all transactions in this period.

¹²For Bitstamp, Coinbase, and Kraken, we use the BTC to USD exchange rate, as discussed above, since it is more likely to be traded by the US investors.

Panel B that the price discrepancies within Europe are much smaller than the overall arbitrage index. The index within Europe, on average, is 1.02. And there are only a few dates, in May 2017 and then December 2017 to January 2018, when the price dispersion is around 6%. And again, a similar picture emerges in Japan; see Fig. 3, Panel C, where the within-region arbitrage index is, on average, less than 1.02 over the same time period. There are two short periods in January and December 2017, where the arbitrage index goes to 1.05. The December increase is parallel to the patterns we observed in the other regions. And finally, in Fig. 3, Panel D we look at Korea. Unfortunately, we only have data for Bithumb starting September of 2017. Therefore, we can only calculate an arbitrage index from that date on. We see that for most of September to December the arbitrage index within Korea was less than 1.03, which is similar to the other regions. But from the end of December to the end of January of 2018, the arbitrage index jumped to 1.05, which is still significantly lower than the full index.

[Fig. 3 About Here]

Overall, the results show that the arbitrage opportunities are much smaller within regions than across regions. Thus, not surprisingly, our results suggest that cryptocurrency exchanges within a given country or region seem to be much better integrated than across regions.

4.3. Price ratio between geographical regions

To confirm that a significant part of the arbitrage spread is driven by price deviations across geographic regions, in Fig. 4 we plot the price ratio at the minute level between the US and each of the other major regions in 2017. Remarkably, we see that for large parts of 2017 to February 2018, prices on Korean exchanges were more than 20% above the US, making Korean exchange prices the highest among all exchanges. There were two distinct periods in June and December to January when the price ratio went as high as 1.6 for a sustained time period. The fact that Korean exchanges have a premium over most of 2017 has even been termed the Kimchi premium. But we also see that Japan had significant price dispersions from the US during the same time periods there was dispersion in Korea. However, the price ratio between Japan and the US had a maximum of 1.2. In contrast, the price differences between the US and Europe are small compared to other regions, which is perhaps not surprising given that the same set of exchanges operate in the US and Europe. The results suggest that a big fraction of the large arbitrage spreads we show for the overall market are driven

by price differences across regions. These differences, in many instances, are persistent over long time periods.

[Fig. 4 About Here]

4.4. *Arbitrage profits*

So far, we have only looked at the maximum and average price dispersion between regions or exchanges. But this analysis does not take into account the volume traded at different prices. To capture the full magnitude of arbitrage opportunities between regions, we now calculate how much profit could have been made with cross-exchange arbitrage. We only look at the period between November 2017 and February 2018 since these were the months that saw the maximum trading liquidity as well as spikes in the arbitrage index.

We calculate the arbitrage profits at the second level and aggregate it at the daily level. To make sure that our results are not driven by price volatility, we only look at seconds where the price difference between the exchanges is larger than 2%.¹³ For each second, we find the aggregate amount of low-priced volume that could have been sold in a high-price region. For that purpose, we calculate the sell-initiated volume in the region that has the lowest price in a given second. By using the sell-initiated volume, we know that we would have been able to buy at this price. Then we calculate the buy-originated volume in the region with the highest price. Again, the same logic holds that this ensures we could have sold in that region. We then compare the two numbers and take the minimum of them. This gives us the volume that could have been traded at a given second between the two regions. We then compute the profit by multiplying this volume by the difference in prices at which it could have been bought and sold. As long as the price difference between the exchanges exceeds 10 %, we allow a more aggressive strategy where all buy-initiated trades in the high price region are executed subject to the constraint that the cumulative outstanding balance between buy and sell trades is less than 25 bitcoins. When the constraint binds or when the price difference becomes lower than 10 %, we only execute sell-initiated trades in the low-price region.

In Fig. 5, Panel A we graph arbitrage profits between the US and Korea. All profits are reported in USD to facilitate comparison between regions. In the beginning of November arbitrage profits on an average day are around \$100,000. However, starting

¹³In case when there are several trades per second, we construct the volume-weighted price per second.

on November 28, the daily profits jump to more than \$2 million a day, and in the middle of December it is up to \$20 million for many days. This level of profits persist from the middle December until the end of January, when the average profit per day is around \$20 million, but there is significant heterogeneity, with several days where the daily arbitrage profit reaches \$40 million a day. The timeline of these profits of course closely follows the arbitrage index that we plotted before. In line with those findings we also see that the arbitrage profits drop off significantly at the end of January 2018. The total profits over the examined four months is \$930 million. For the exchanges that are not in the Kaiko data, we do not have signed volume. In the case of Korea this means we only have unsigned volume data for Korbit. Assuming that the buy- and sell-initiated volume dynamics are similar on Bithumb and Korbit, one can get a more realistic number of total arbitrage profits in Korea by scaling the profits by the total volume on Korean exchanges. For that purpose, we compute the ratio of total volume on Korbit and Bithumb to the volume on Bithump and multiply the profits by this ratio. When using this adjustment, the arbitrage profits between Korea and the US over the four-month period become \$1.275 billion.

[Fig. 5 About Here]

We repeat the same exercise for the arbitrage profits between Japan versus the US in Fig. 5, Panel B. As before we, see that the level of the arbitrage profits jumps up at the beginning of December 2017 to almost \$20 million a day. It then falls back and again increases to above \$25 million for several days at the end of December. In January the average daily profits are about \$7 million, with several days higher than \$30 million. Consistent with the arbitrage index, the arbitrage profits become small at the end January. The total profits over the examined four months is \$322 million. Again, we can adjust the profits by the volume of the exchanges for which we do not have signed volume as above. For Japan, these exchanges are Coincheck and Fisco. Once we include the volume on these exchanges, the total profit over the four months increases to \$675 million.

Finally, in Fig. 5, Panel C, we also show the arbitrage profits between Europe and the US. As expected, the profits are much smaller between these two regions since the prices seem more integrated. We show that the maximum daily arbitrage profits are around \$2 million to \$3 million, but for the modal day in December and January they are around \$500,000. The total profits over the examined four months is \$25 million.

5. Correlation structure of price deviations across regions

We now analyze the correlation structure of price deviations across different regions. Just by eye-balling the price ratios between Korea and the US and Japan and the US in Fig. 4 above, there appears to be a strong inter-temporal correlation when arbitrage spreads open up and close. To study the correlation structure more formally, we use the extended sample of exchanges that cover 19 regions. For each region, we form the ratio of the volume-weighted average price across the exchanges in a given region to the volume-weighted average bitcoin price in the US at the minute level. We use the US price as a proxy for the world market price; we get the same results when we use the average-weighted price across all large exchanges.

We report a correlation matrix of the price ratios between countries in Table A2 in the appendix. To visualize the correlation structure of arbitrage spreads, we plot a heat map in Fig. 6 where lighter colors signify higher correlations. We can see that the correlation between arbitrage spreads, on average, is very high; more than half of the countries show a correlation greater than 50%. And in some regions, the correlation is well above 75%; see, for example, Indonesia, Australia, Singapore, Japan, or South Korea. This result is surprising since we would not have expected that these deviations necessarily move in lockstep. While limits to arbitrage can explain the existence of price deviations, these theories do not make any prediction about them being positively correlated. We will discuss the implications of our findings in more detail in Section 8.

[Fig. 6 About Here]

5.1. *Co-movement of price deviations and buying pressure*

To better understand what might be driving the correlations between price ratios, we form a measure of buying pressure in the US. The idea is to identify time periods of a particularly quick appreciation of bitcoin prices. Since we show later that bitcoin prices react strongly to order flows, these periods also coincide with the times when there is a particularly strong increase in demand for bitcoin worldwide. To estimate this buying pressure, we use the standard Hodrick-Prescott filter to form a series of the smoothed log bitcoin price at the weekly level in the US. We then calculate the deviations of the actual log bitcoin price from the smoothed log price. Fig. 7, Panel A plots the time series of the smoothed log bitcoin priced using the Hodrick-Prescott filter, and Panel

B plots the residuals between the actual and the smoothed log price, our measure of buying pressure. We see that the residual deviations are particularly large in January, May, and October 2017 and then in December and beginning of January 2018. This is in line with the general perception of when bitcoin market was heating up.

[Fig. 7 About Here]

We then regress the deviations of a country's bitcoin price relative to the US on our measure of buying pressure. For each country, this provides us with a measure of the sensitivity of this country's arbitrage spread to the buying pressure factor; we call this the arbitrage beta. The results are reported in Table 5. On average, we find a strong positive beta: arbitrage spreads outside the US open up when there is an increase in buying pressure in the US. Thus the buying pressure factor explains the strong positive correlation between arbitrage spreads we report in Table A2.

In a final step, we show that the countries that, on average, have a higher premium over the US bitcoin price are also those with a higher arbitrage beta. For this purpose, we run a cross-sectional regression of the average bitcoin price premium in each country relative to the bitcoin price in the US (calculated from January 2017 to February 2018) on this country's arbitrage beta. Fig. 8 shows a strongly positive and significant slope. So countries that, on average, have a higher bitcoin premium relative to the US also respond more strongly by widening arbitrage deviations in times when buying pressure goes up in the US.

[Fig. 8 About Here]

Our results thus show that the marginal investor outside the US and Europe is willing to pay more for bitcoin in response to positive news or sentiments. How can one explain this difference in valuation? We conjecture that they might reflect weaker financial institutions or tighter capital controls in the countries outside the US and Europe. The marginal investor in a country with poorly functioning financial institutions or tighter capital controls might be willing to pay more for bitcoin since they would benefit more from the adoption of cryptocurrencies. Therefore, any news about the potential adoption of Bitcoin (or any sentiment change) would increase the price in these countries more.

Of course, the price differences could only persist if capital markets are segmented or capital is slow to flow across borders. Controls on the flow of fiat currency between regions reduce the efficient use of arbitrage capital. Arbitrage profits are realized by buying bitcoins in regions with low bitcoin prices, say the US, and selling in regions with high bitcoin prices, say Korea. This trade requires capital in the US and generates

profits in Korea. If capital cannot be repatriated seamlessly from Korea to the US, arbitrage capital can become stuck within a country and thus become scarce.

5.2. *Co-movement of price deviations and capital controls*

To test the importance of capital controls on fiat currencies, we analyze whether the positive correlation in arbitrage spreads between countries we show above is explained by the level of openness of a country. If countries that are relatively closed have a higher convenience yield for bitcoin, then we should see their arbitrage spreads (relative to the world market price) move more closely together. Countries that are more open should not be correlated since any price deviation will immediately be arbitrated away. To measure the level of openness of a country, we use a capital control index developed by [Fernandez et al. \(2015\)](#).¹⁴ As a measure of the pairwise tightness in capital controls between two countries, we take the product of the capital control index of the two countries. This measure is zero if at least one of the countries is totally open and approaches one if both countries have very high levels of capital controls. It is important to note that this measure depends on the match between two countries. For example, the pairwise capital control measure between Turkey and South Africa is quite high since both countries have significant capital controls. However, at the same time, the pairwise capital control between Turkey and the United Kingdom is low since at least one of the countries (UK) has an open capital market.

In [Fig. 9](#), we regress the correlation in arbitrage spreads between two countries on our measure of pairwise capital controls. The slope of this regression is 0.29 and statistically significant at the 1% level. In support of the idea that capital market segmentation is important in explaining arbitrage spreads in bitcoin prices, we find that there is a significantly positive relation between the correlation of arbitrage spreads and capital controls. In other words, two countries that are both relatively closed to capital flows have a higher correlation in arbitrage spreads. However, if at least one of the countries has a more open capital market, which means arbitrage capital can flow more freely, the correlation goes down significantly.

[[Fig. 9](#) About Here]

¹⁴We use the aggregate index “ka” of the intensity of controls. See [Fernandez et al. \(2015\)](#) for more details.

6. Arbitrage in other cryptocurrency markets

To further investigate if controls on fiat currencies play an important role in explaining arbitrage spreads, in this section we analyze the price of ethereum and ripple relative to fiat currencies and then relative to bitcoin. If capital controls on fiat currencies play an important role, we would expect to find large arbitrage spreads between ethereum to fiat currencies but much smaller arbitrage deviations between two cryptocurrencies.

Fig. 10 plots the arbitrage index for ethereum and ripple. We can see that, similar to the bitcoin arbitrage index, there is significant variability during the year. Periods of relatively low levels of the arbitrage index alternate with prolonged spikes. Similar to bitcoin, at the height of its peak December and January, the ethereum arbitrage index stays at about 1.5. The ripple index displays similar behavior, but the series only starts from August 2017 due to the data availability.

[Fig. 10 About Here]

As in the case of bitcoin, a significant part of the arbitrage spread in ethereum and ripple prices is driven by price deviations across geographic regions. We do not report this analysis in the paper since they are very similar to the reported arbitrage indices for bitcoin, but they can be obtained from the authors on request. Furthermore, by comparing the three arbitrage indices, one can notice the high degree of correlation between them. All three arbitrage indices usually spike at about the same time and take similar levels.

6.1. Arbitrage between cryptocurrencies

To analyze if the same arbitrage spreads exist between cryptocurrencies, we focus on ethereum as the second-most traded cryptocurrency after bitcoin.¹⁵ We only look at the months of November 2017 to February 2018 since these are the time periods when the BTC price has the strongest price dispersion relative to fiat currencies. It is also the time period when trading in ethereum and other coins become more liquid. If constraints in the movement of capital contributes to the arbitrage profits between BTC and the local fiat currencies, then these price deviations should be much smaller across cryptocurrencies, which, by design, do not obey the same restrictions.

¹⁵We run the same analysis for ripple and obtain qualitatively similar results, which are available upon request from the authors.

In Fig. 11 we plot the ratios of the value weighted price of bitcoin to ethereum at the minute level across different regions. This is very similar to the calculations in Section 4.3. However, here we make two modifications to the process. First, we calculate the volume-weighted price at the five-second level since trading in ethereum has less volume than bitcoin. Second, since not all exchanges directly provide a platform to trade bitcoin to ethereum, but often trade bitcoin and ethereum only to the local fiat currency, we calculate the local exchange rate of bitcoin to ethereum as the cross-rate. For example, to calculate the exchange rate of ethereum to bitcoin on Bithumb, we take the ratio of the exchange rate of ethereum to Korean won to the exchange rate of bitcoin to Korean won.

[Fig. 11 About Here]

Fig. 11, Panel A shows the ratio of the two exchange rates of bitcoin to ethereum between the US and Japanese exchanges from November 1, 2017, to February 28, 2018. As mentioned before, if there were no frictions in the currency markets, then this ratio should be constant and equal to one all the time. We see that the price ratio indeed lies in a narrow band between 0.98 and 1.02.

We repeat the same calculations in Fig. 11, Panel B for the price of bitcoin to ethereum, but for the US and Korea. Again, we see that the deviations from the ratio of one are relatively small, and even in December and January it hovers around 1.03 to 0.97. In comparison, these were the months when the Kimchi premium (price of bitcoin in fiat currency in Korea versus US) was highest, almost 50% for several days. This again confirms that the arbitrage opportunities are much less pronounced and persistent between different cryptocurrency markets than between cryptocurrency and fiat currency markets. We finally repeat the same exercise in Fig. 11, Panel C for the US and Europe and find that the differences in the price of BTC to ethereum across all three months are small—less than 1% for the average day—and there are only several days in mid-December where the ratio is around 1.03. However, this lack of price dispersion should not be too surprising since we have previously shown that, even in the bitcoin to fiat currency market, the difference between US and European exchanges is smaller than in other regions.

7. Order flow and prices

To provide an estimate of how much capital is required to close the arbitrage spreads we show above, we develop and estimate a model of order flows and prices. The existing literature shows the importance of net order flows for price formation in traditional

financial markets.¹⁶ While previous research attributes the price pressure of net order flow to price discovery, it is less clear what the fundamentals are in the case of cryptocurrency markets and whether there are any traders who have more information than others. Nevertheless, in this section we show that a strong positive relation between net order flow and prices also exists in cryptocurrency markets.

A common way to estimate the impact of net order flow is to regress price differences or returns over a particular time period on the signed volume of trades during the same period. The complication in the bitcoin market is that the same asset is traded simultaneously on multiple exchanges and, as we showed earlier, often at different prices. Therefore, when forming their demand, investors might not only look at prices on their own exchange but also take into account prices on the other exchanges where bitcoin is traded. As a result, a regression of returns on signed volume in each market separately may give a biased picture of the true impact of net order flow.

To accommodate the case of multiple exchanges, we decompose signed volume on each exchange into a common component and an idiosyncratic, exchange-specific component:

$$s_{it} = \bar{s}_i + \beta_i^s s_t^* + \hat{s}_{it}, \quad (1)$$

$$E[s_t^*] = 0, \quad E[\hat{s}_{it}] = 0, \quad E[s_t^* \hat{s}_{it}] = 0.$$

Here s_{it} is signed volume on exchange i , s_t^* is the common component for all exchanges, \hat{s}_{it} is an exchange-specific component, and \bar{s}_i is the exchange-specific mean. Similarly, we decompose the log return on each exchange, $r_{it} = \ln(p_{it}/p_{it-1})$, into a common component and an idiosyncratic, exchange-specific component:

$$r_{it} = \bar{r}_i + \beta_i^r r_t^* + \hat{r}_{it}, \quad (2)$$

$$E[r_t^*] = 0, \quad E[r_{it}] = 0, \quad E[r_t^* \hat{r}_{it}] = 0.$$

The models (1) and (2) can be estimated either separately by factor analysis by assuming additionally that

$$E[\hat{s}_{it} \hat{s}_{jt}] = 0, \quad E[\hat{r}_{it} \hat{r}_{jt}] = 0, \quad \text{for } i \neq j,$$

or jointly by the canonical correlation analysis; see [Jolliffe \(2002\)](#) for a textbook treat-

¹⁶See, for example, [Evans and Lyons \(2002\)](#), [Berger et al. \(2008\)](#), and [Fourel et al. \(2015\)](#) for the study of foreign exchange markets; [Brandt and Kavajecz \(2004\)](#) for US Treasury markets; [Deuskar and Johnson \(2011\)](#) for the S&P 500 futures market; and [Chordia, Roll, and Subrahmanyam \(2002\)](#), [Goyenko, Holden, and Trzcinka \(2009\)](#), and [Hendershott and Menkveld \(2014\)](#) for NYSE stocks.

ment. The canonical correlation analysis estimates models (1) and (2) by maximizing the correlation between s_t^* and r_t^* . The factor analysis and the canonical correlation analysis are linear models. They both estimate the common factors as a linear combination of input data; that is,

$$s_t^* = \sum_i w_i^s (s_{it} - \bar{s}_i), \quad (3)$$

$$r_t^* = \sum_i w_i^r (r_{it} - \bar{r}_i), \quad (4)$$

where w_i^s and w_i^r are called the factor weights. Both models (1) and (2) imply that

$$\sum_i w_i^s \beta_i^s = 1, \quad \sum_i w_i^r \beta_i^r = 1.$$

We fix the scale of the common factor in signed volume by requiring that the sum of factor loadings, β_i^s , is equal to one. To fix the scale of the common factor in returns, we require that the sum of factor weights, w_i^r , is equal to one. Under this normalization the common factor in returns becomes a portfolio. Suppose one buys β_i^s bitcoins on each exchange i . Since both the sum of β_i^s and $w_i^s \beta_i^s$ is equal to one, the total amount of bitcoins bought is equal to one, and the common component in signed volume is increased by one. Hence, in the regression

$$r_t^* = \lambda s_t^* + \varepsilon_t, \quad (5)$$

the coefficient λ measures the price pressure of the aggregate order flow.

While the price of bitcoin across exchanges can be different for some period of time, as we show in the analysis above, one should expect the price of bitcoin across any two exchanges to be cointegrated. More generally, any linear combinations of prices where the sum of weights is equal to one should be cointegrated as well. Thus, the restriction that the sum of factor weights, w_i^r , is equal to one allows us to decompose the price on each exchange into a common component and an exchange-specific deviation from the common component:

$$p_{it} = p_t^* + \hat{p}_{it}, \quad p_t^* = \sum_i w_i^r p_{it}. \quad (6)$$

Unlike the common component of the price, p_t^* , each \hat{p}_{it} must be a bounded process. Since we use the log-prices, \hat{p}_{it} measures the percentage deviation from the weighted average price across exchanges. If any of \hat{p}_{it} were an unbounded process, it would imply arbitrarily large arbitrage opportunities.

7.1. *Decomposition of returns and order flow*

We estimate the models (1) and (2) using data from 14 exchanges at the five-minute, hourly, and daily frequency. We exclude Binance and Chinese exchanges from this analysis because the Binance data are only available starting from November 2017, and the Chinese exchanges stopped trading in September 2017.

The results of this estimation are reported in Tables 6 and 7. The first panel of Table 6 reports the factor loadings, weights, and R -squared of the factor analysis of signed volume using data at the five-minute frequency. Not surprisingly, the factor loadings are highest for the exchanges with the largest volume. Bitfinex has the highest loading of 0.35 followed by Coinbase USD, Bitstamp USD, and bitFlyer. With the exception of Quoine, the common component explains from 30% to 60% of the variation in exchange-specific signed volume. In the next two panels of Table 6, we repeat the same analysis but at an hourly and daily frequency. The results show that at longer frequencies, the common component of signed volume explains an even higher fraction of variation; R -squares go up to 70%. The coefficients on the factor loadings stay relatively stable, and the volume on the exchanges, which are less liquid and less integrated with the rest of the markets, also load less heavily on the common component of signed volume at the longer frequencies.

In Table 7 we repeat a similar analysis for the common component of returns. As discussed above, the one difference is that we now normalize the weights across the exchanges to sum up to one. We start with the five-minute frequency of the return data. For returns the common component is even more important than for signed volume. Even at the five-minute frequency, the common component in returns explains about 80% of exchange returns on exchanges where the US dollar and tether are a base currency. The only exception is Kraken, where the R -squared is only 40%. The R -squared is also low for Japanese and Korean exchanges. As shown in the next two panels of Table 7, the same pattern persists at longer frequencies. But at longer frequencies the common component becomes progressively more important. We find that the common component in returns explains around 90% in variation at the hourly frequency and 96% at the daily frequency. This is intuitive since we have shown above that the average correlation in returns at the hourly and daily level are extremely high, upside of 95%.

We also estimate models (1) and (2) using the canonical correlation analysis. In all cases, the extracted factors are very similar to those estimated using the factor analysis. At the five-minute and hourly frequency, the correlation between factors is above 98%. At the daily frequency the common factors in returns correlate at 98%

with each other, and correlation between the common factors in signed volume is 90%. Therefore, in what follows, we only report results based on the factors extracted using the factor analysis.

7.2. *Common components of returns and order flow*

In Table 8 we show that common component in signed volume explains a very large fraction of the common component in returns. The first three columns report the estimates at the five-minute frequency; the next six columns show results at the hourly and then daily frequencies. In column (1) we regress the common component in returns on the contemporaneous common component in signed volume. The coefficient on the independent variable is 8.8×10^{-6} with a t-statistic of 80 and an R -squared of 54%, which shows that there is a very strong relation between the common component in returns and signed volume. For any 10,000 bitcoin increase in buy volume across exchanges, we see a 9% increase in the price on average. Assuming that the price of bitcoin is \$10,000, this translates into the price impact of nine basis points for a \$1 million trade. The price impact is considerably larger than that observed in foreign exchange markets, where according to [Berger et al. \(2008\)](#), the price impact is 5.4 basis points and 7 basis points in euro-dollar and dollar-yen markets for a \$100 million trade. But the price impact in the bitcoin market is considerably smaller than that in the US stock market. [Goyenko, Holden, and Trzcinka \(2009\)](#), following the methodology developed in [Hasbrouck \(2009\)](#), show that the average price impact across randomly chosen 400 NYSE stocks is 16 basis points for a \$10,000 trade.

In columns (2) and (3) we look at the persistence of the price impact. In column (2) we add one lag of the common component in signed volume, and in column (3) we add five lags. We see that the coefficients at all five lags are negative and significant, which suggests that part of the price pressure in the common component is temporary. A bit less than half of the impact on returns reverses within the next five periods. The negative sign is also consistent with the negative sign of the first-order correlation of five-minute returns. Note that the reversal at the five-minute frequency persists for longer than just five lags. In unreported regressions we repeat the estimation with 12 lags, and still find significant reversal for all the lags (i.e., negative and significant coefficients on all lag coefficients). However, to conserve the space, we only report five lags in all cases.

In the next three columns we report the results of the price impact regressions at the hourly level. The price impact of the contemporaneous common factor in signed volume is still positive and highly significant but slightly smaller than that for the five-

minute frequency. Buying 10,000 bitcoins over an hour predicts a 6% increase in the price on average. Similar to the five-minute frequency, the coefficients on lagged signed volume are negative, suggesting that even at the hourly frequency some of the price impact is temporary. Finally, at the daily frequency we see a smaller price impact on average (the coefficient on the contemporaneous common component in signed volume is only 3.6) and also much reduced mean reversion in lags two through five. Buying 10,000 bitcoins over a day predicts a 3.6% increase in the price on average. The past signed volume is only significant at the first lag, bringing the total price impact to about 3% for every 10,000 bitcoins.

We also repeat our analysis of the price impact separately for the first and second half of our sample. We find that the estimated coefficient on the price impact is qualitatively the same in both subsamples. This result confirms that our estimate is stable over our sample period and not driven by any short term trends.

Overall, we show that the common component of signed volume explains a large fraction of the common component of returns at all frequencies, and subsequent mean reversion is much smaller at the daily level than at higher frequencies.

7.3. *Idiosyncratic price pressure*

To find the exchange-specific price pressure we follow [Hasbrouck \(1991\)](#) and estimate the Vector Autoregression (VAR) model of the idiosyncratic part of signed volume and the exchange-specific deviation from the common component:

$$\hat{s}_{it} = \sum_{s=1}^{\tau} b_{i,s} \hat{s}_{it-s} + \gamma_i \hat{p}_{it-1} + u_{it}, \quad (7)$$

$$\hat{p}_{it} = \sum_{s=1}^{\tau} a_{i,s} \hat{p}_{it-s} + \lambda_i \hat{s}_{it} + v_{it}, \quad (8)$$

where

$$E[v_{it}] = E[u_{it}] = 0, \quad E[v_{it}v_{is}] = E[u_{it}u_{is}] = 0, \text{ for } s \neq t, \quad E[v_{it}u_{is}] = 0.$$

We estimate these equations as a system of Ordinary Least Squares (OLS) equations and obtain the results for Eq. (7) in [Table 9](#) and for Eq. (8) in [Table 10](#). The idiosyncratic component of the price on each exchange is obtained as the residual value after taking out the common component from each price. The exchange-specific deviations from the common component are then estimated as a function of contemporaneous signed volume as well as past deviations from the common price; we use three lags here.

We do not include lagged values of idiosyncratic signed volume since our estimates of Eq. (7) show that the lagged values have low correlation with the contemporaneous value of residual of signed volume. To conserve space, we report the results based on the five-minute frequency.

Each column of Table 10 reports the result of Eq. (8) for each exchange as labeled on the top of the column. By looking at the coefficients on the lagged values of \hat{p}_{it} , it is apparent that \hat{p}_{it} are highly persistent mean-reverting processes. All coefficients on the past three lags are positive. The sum of the three lags, while below one, is close to one for all exchanges. When the price on any exchange deviates above (below) from the average price on other exchanges, the subsequent returns on this exchange are predicted to be lower (higher) than the returns on other exchanges. But the convergence to the common component is slow. The coefficient on the past lags of \hat{p}_{it} are particularly high for Japanese and Korean exchanges, such as Bitflyer or Bithumb. This confirms our prior results that arbitrage spreads persist for longer on these exchanges.

Compared to the price pressure we estimate for the common component, the idiosyncratic (exchange-specific) price pressure is significantly higher on almost all exchanges. The price pressure is particularly high on smaller and less liquid exchanges, for example, Zaif or Coinbase EUR. Bitfinex and Bitstamp are the two exchanges where the estimated coefficient on the idiosyncratic price impact are lower than the one on the common component. These are two of the largest and most liquid exchanges. Note that one should be careful with the interpretation of the exchange-specific price pressure. When prices on one exchange are either very high or low, traders might adjust which exchanges they trade and how they trade. Hence, the idiosyncratic part of signed volume might change endogenously, and this can lead to non-linearity in the relation between price and signed volume, which is not picked up in our model. Nevertheless, our results show that the exchange-specific part of signed volume plays an important role at explaining the deviation of prices on an exchange from the common component.

8. Discussion of arbitrages and constraints

8.1. *Implementation of arbitrage strategies*

Our goal in this section is to outline the mechanics of the different arbitrage strategies that exist in the bitcoin market and the potential risks and costs that can impede the effectiveness of arbitrage. Consider the situation that the price in Korea is above the price in the US as we show above. In the world without frictions, this situation would constitute a riskless arbitrage. One could buy bitcoins in the US, sell them

for Korean won in Korea, exchange won for dollars, and then transfer dollars back to the US. In practice, this textbook arbitrage is not possible since the nature of bitcoin transactions implies that it takes about an hour for the transaction to be registered on the bitcoin blockchain. Moreover, exchanges typically take from a few hours to several days to transfer fiat currency. In that time period the arbitrage opportunity might disappear. As a result, to lock in the arbitrage, an arbitrageur has to simultaneously buy bitcoin on the exchange where the price is low and sell it on the exchange where the price is high.

Ideally, the arbitrageur would like to short sell bitcoin on the market where the price is high, say Korea, and buy bitcoin in the US. Then she would transfer bitcoin from the US to Korea and so realize the risk-free profit. This trading strategy, however, is not always feasible because only a few exchanges allow short sales. Some of the larger exchanges such as Kraken, Bitmex, and GDAX offer short support, especially to their larger clients. A few exchanges such as Bitfinex even allow taking leveraged short positions. However, exchanges in Korea and Japan, which for extended time periods were trading often more than 10%-25% above other exchanges, do not allow for short selling.

In the absence of short sales on a particular exchange, the arbitrageur can resort to two alternative arbitrage strategies. First, she could establish a negative position in bitcoin by trading on margin, which is similar to short sales but does not allow for physical settlement. In this case, the arbitrageur can profit from the trade only if prices on the two exchanges converge in the future. Thus, the arbitrageur is subject to the convergence risk, which has been extensively studied in the limits of arbitrage literature; see, for example, [Shleifer and Vishny \(1997\)](#) or [Gromb and Vayanos \(2002\)](#). While, in theory, prices across exchanges may not converge for a long time, [Fig. 4](#) shows that, in practice, arbitrage opportunities in any given market open for less than two days, on average, and even in the extreme never existed for more than a month.

The second arbitrage strategy is to hold a positive balance of bitcoins on both exchanges and simultaneously buy and sell bitcoins across the two exchanges whenever the price on one exchange deviates from that on the other. Naturally, the bitcoin balance of the arbitrageur will go down on the exchange where the price of bitcoin is high (since this is where she would sell bitcoin) and increase on the exchange where the price is low. To replenish it, the arbitrageur needs to transfer bitcoins from the exchange with high bitcoin balance to the one with low balance and vice versa for capital. In an ideal world, she would like to instantaneously shift trading gains from exchanges where she sold bitcoin to the ones where bitcoin is cheap and then repeat the arbitrage transaction. Arbitrage becomes more efficient the quicker the arbitrageur can

recycle capital from one account to another, but in some exchanges this transfer can take up to several days. We discuss below some of the constraints to capital movements that reduce the speed of transferring capital across exchanges. While this strategy does not expose the arbitrageur to convergence risk, a drawback of this strategy is that the arbitrageur becomes exposed to bitcoin price fluctuations. To mitigate this risk she can establish short positions either on some of the exchanges described above or by borrowing bitcoin from people who hold large amounts of bitcoins without an interest to sell, the so-called *hodlers*.¹⁷ Of course, these hodlers themselves would be in a great position to do the arbitrage in cryptocurrency markets. Starting from the end of December 2017, the arbitrageur can also use Chicago Board Options Exchange (CBOE) and Chicago Mercantile Exchange (CME) bitcoin futures contracts to hedge the price risk. The futures contracts track bitcoin price, on major US dollar exchanges and have an average daily open interest of about 10,000 bitcoins.¹⁸

8.2. *Constraints to arbitrage*

In practice, the arbitrageur has to incur a number of transaction costs, but their magnitudes are too small to prevent arbitrageurs from implementing the above trading strategies. To transfer bitcoins, the transaction has to be recorded on the Bitcoin blockchain; this is the work of the so-called miners that provide certification of transactions and add blocks to the blockchain if they win the hashing competition. The fees peaked around \$40 in the end of December 2017 at the height of the bitcoin price but since February have come down to below \$10. Since these are fixed cost, they are minuscule relative to the size of the potential arbitrage. In addition, exchanges have trading fees, which increase the cost of trading. In the appendix we show the magnitude of the fees for the exchanges used in this paper. These fees range from 0.25% of the amount traded to 0.1%. Most exchanges do not charge fees on a trade by trade basis but assign them based on the trading volume in a given month or week. Furthermore, most exchanges charge zero fees for trades that add to the liquidity of the order book. The exchange fees are comparable to the bid-ask spreads, which are, on average, between 1 and 10 basis points. Finally, many exchanges charge withdrawal

¹⁷The term hodler is a peculiarity of the bitcoin market since one investor in bitcoin wrote in a post on the bitcoin talk forum in 2013 while prices were dropping “*I AM HODLING*”. This has become a meme for “Hold On for Dear Life”.

¹⁸The CBOE contract settlement price is determined by results of the auction on the Gemini exchange. The CME contract settlement price is based on the CME CF bitcoin Reference Rate (BRR), which aggregates the price from major US dollar exchanges (see <http://www.cmegroup.com/trading/equity-index/us-index/bitcoin.html> for more details.)

fees; these range from 10 to 50 basis points per withdrawal for most of the exchanges. But all large exchanges state that for large traders they provide preferential customized fees that are far below the cost for retail investors. In sum, these fees are small for large transactions. Overall, we believe that for large players the round-up trading costs should be within 50 to 75 basis points. These transaction costs are very low compared to the arbitrage spreads we show, and therefore cannot explain the arbitrage spreads we find in the paper.

Another factor that might limit the willingness of traders to engage in arbitrage is the governance risk of cryptocurrency exchanges. The governance risk of exchanges arises since, in practice, to trade on an exchange, the arbitrageur has to transfer her bitcoins to the exchange and therefore give up control of her coins to the exchange. Judging from many widely publicized hacks of exchanges, these can lead to significant losses to investors who trade there.¹⁹ However, it seems unlikely that this explains the arbitrage spreads we found. Concerns about the governance risk of an exchange should affect its volume and possibly bid-ask spreads. But we show that many of the exchanges with the largest arbitrage spreads, for example Bithumb and Korbit, have very significant volumes and small bid-ask spreads. Moreover, we show that arbitrage spreads are much larger across, rather than within, regions. For exchange risk to explain this pattern, one would have to assume that it is correlated within a region. But this is not supported by our data since there is significant heterogeneity in the liquidity of exchanges within a region, but nevertheless arbitrage spreads are small between them. While governance risk does not seem to explain the direction and size of the arbitrage spreads across exchanges, it might well explain why many institutions might not want to participate in this market altogether.

Finally, an important potential constraint to arbitrage are cross-border capital controls. As we described before, unless the arbitrageur is willing to bet on price convergence between Korean and the US exchanges, she would need to sell bitcoin in Korea and repatriate profits from Korea to the US. The regulation in some countries make cross-border transactions difficult for retail investors. As referenced before, in Korea, local residents and companies moving more than \$50,000 out of the country in a single year must submit documents to authorities proving their reasons for the transfers, which may not always be approved. Industry reports, as well as descriptions from trading blocks, suggest that these constraints are binding for retail investors. However, it is more difficult to quantify how binding these constraints are for large financial

¹⁹For example, in the notorious hack of Mt Gox in 2014 650,000 bitcoins were stolen from customers and the company.

institutions that trade in multiple international financial markets.²⁰ There are a few reports that suggest large institutions are able to avoid these constraints. In a recent IMF working paper, [Baba and Kokenyne \(2011\)](#) find that the effectiveness of capital controls in South Korea seems limited since capital flows in and out of the country and the effectiveness of monetary policy does not seem to be significantly changed after the introduction of capital controls in the early 2000s. Similarly, industry reports suggest that there are networks of forex dealers that help institutions to transfer capital in and out of the country. Thus, capital controls should not impose insurmountable constraints to arbitrage across regions, especially for large traders, but they add to the cost of arbitrage. This interpretation is supported by our finding that arbitrage spreads are an order of magnitude smaller in two-way cryptocurrency trades (say bitcoin to ethereum) on the exact same exchanges where we see big (and persistent) arbitrage spreads relative to fiat currencies. But even in case of the fiat currency, the arbitrage spread does not stay open for more than a month and eventually closes.

In summary, our analysis suggests that the history of bitcoin exchanges over the past two years was marked by recurring episodes of arbitrage opportunities opening up and closing again and a few periods of extremely large arbitrage spreads that persisted for several weeks. Most of the time, arbitrageurs are able to equalize prices across markets.²¹ But at times the arbitrage capital seems to get overwhelmed by the noise traders who are driving up the price in certain markets or lose heart when negative information about bitcoin comes out. We also show above that arbitrage spreads within regions are correlated with periods when arbitrage spreads across countries are also particularly large. This pattern is again consistent with the idea that arbitrage capital is limited in the short run, so when arbitrage capital is allocated to areas where the arbitrage profits are particularly high, opportunities open up in other places.

9. Conclusion

This paper studies arbitrage and price formation in the cryptocurrency market. We show that there are large and recurring deviations in cryptocurrency prices across exchanges that open up across different exchanges and often persist for several days and weeks, pointing to significant market segmentation. The arbitrage spreads are much

²⁰A related constraint is that many retail investors face restrictions on which exchanges they can trade. For example, foreign nationals are typically prevented from opening up accounts and trading on local exchanges. But similar to capital controls, large financial institutions should be able to bypass these restrictions and operate across regions.

²¹Industry reports suggest that hedge funds and high frequency traders have been active across different cryptocurrency markets for several years.

larger for exchanges across different countries than within the same country. In contrast, exchange rates between different cryptocurrencies show much smaller deviations on all exchanges. These results suggest that capital controls, together with the lack of regulatory oversight on crypto exchanges, are the main factors contributing to market segmentation.

Because of the capital controls arbitrageurs find it difficult to scale up their trading strategies with the intensity of noise-trader activity in a timely fashion. While regulations in some countries make cross-border transfers in fiat currencies difficult for retail investors, large institutions are typically able to avoid these constraints. However, the lack of regulatory oversight may create impediments for large public institution to enter the cryptocurrency space and slow down the supply of arbitrage capital.

This market segmentation allows us to measure differences in pricing behavior across markets. We show that there is significant co-movement of arbitrage spreads across countries. The spreads in the rest of the world are almost always positive relative to the US and Europe and go up more during times of large bitcoin appreciation. The correlation in arbitrage spreads between countries is stronger for countries that have stricter capital controls than more open economies. This pattern suggests that the marginal investors who price cryptocurrencies in countries with less developed capital markets value cryptocurrencies more highly, possibly because they have a higher convenience yield for bitcoin.

To examine how these price deviations between exchanges emerge, we analyze the relation between net order flows and prices in the cryptocurrency market. We decompose signed volume and returns on each exchange into a common component and an idiosyncratic, exchange-specific component. The common component of signed volume explains about 50% of the variation in returns at the five-minute and hourly level, and up to 85% at the daily level. The exchange-specific residuals of signed volume explain variation in exchange-specific residuals of returns at the five-minute and hourly level. We also show that when the price on any exchange deviates above (below) from the average price on other exchanges, subsequent returns on this exchange are predicted to be lower (higher) than the returns on other exchanges.

References

- Athey, S., Parashkevov, I., Sarukkai, V., and Xia, J., 2016. Bitcoin pricing, adoption, and usage: theory and evidence. Unpublished working paper. Stanford University.
- Baba, C., Kokenyne, A., 2011. Effectiveness of capital controls in selected emerging markets in the 2000s. Unpublished working paper. IMF.
- Bekaert, G., Harvey, C. R., Lundblad, C. T., Siegel, S., 2011. What segments equity markets? *Review of Financial Studies* 24, 3841–3890.
- Berger, D., Chaboud, A., Chernenko, S., Howorka, E., Wright, J., 2008. Order flow and exchange rate dynamics in electronic brokerage system data. *Journal of International Economics* 75, 93–109.
- Bohme, R., Christin, N., Edelman, B., Moore, T., 2015. Bitcoin: economics, technology, and governance. *Journal of Economic Perspectives* 29, 213–238.
- Brandt, M. W., Kavajecz, K. A., 2004. Price discovery in the US Treasury market: the impact of order flow and liquidity on the yield curve. *Journal of Finance* 59, 2623–2654.
- Budish, E., Cramton, P., Shim, J., 2015. The high-frequency trading arms race: frequent batch auctions as a market design response. *Quarterly Journal of Economics* 130, 1547–1621.
- Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. *Journal of Financial Economics* 65, 111–130.
- Ciaian, P., Rajcaniova, M., Kancs, D., 2016. The economics of bitcoin price formation. *Applied Economics* 48, 1799–1815.
- Cong, L. W., He, Z., Li, J., 2019. Decentralized mining in centralized pools. Unpublished working paper. NBER.
- DeLong, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98, 703–738.
- Deuskar, P., Johnson, T. C., 2011. Market liquidity and flow-driven risk. *Review of Financial Studies* 24, 721–753.
- Easley, D., O’Hara, M., Basu, S., 2017. From mining to markets: the evolution of bitcoin transaction fees. Unpublished working paper. Cornell University.

- Evans, M. D. D., Lyons, R. K., 2002. Order flow and exchange rate dynamics. *Journal of Political Economy* 110, 170–180.
- Fernández, A., Klein, M. W., Rebucci, A., Schindler, M., Uribe, M., 2015. Capital controls measures: a new dataset. Unpublished working paper. NBER.
- Fourel, V., Rime, D., Sarno, L., Schmeling, M., Verdelhan, A., 2015. Common factors, order flows, and exchange rate dynamics. Unpublished working paper. MIT, City University London, BI Norwegian Business School, and Goethe University Frankfurt.
- Froot, K. A., Dabora, E. M., 1999. How are stock prices affected by the location of trade? *Journal of Financial Economics* 53, 189–216.
- Goyenko, R. Y., Holden, C. W., Trzcinka, C. A., 2009. Do liquidity measures measure liquidity? *Journal of Financial Economics* 92, 153–181.
- Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics* 66, 361–407.
- Gromb, D., Vayanos, D., 2018. The dynamics of financially constrained arbitrage. *Journal of Finance* 73, 1713–1750.
- Harvey, C. R., 2016. Cryptofinance. Unpublished working paper. Duke University.
- Hasbrouck, J., 1991. Measuring the information content of stock trades. *Journal of Finance* 46, 179–207.
- Hasbrouck, J., 1995. Many markets: determining the contributions to price discovery. *Journal of Finance* 50, 1175–1199.
- Hasbrouck, J., 2009. Trading costs and returns for US equities: estimating effective costs from daily data. *Journal of Finance* 54, 1445–1477.
- Hasbrouck, J., Seppi, D. J., 2001. Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59, 383–411.
- Hendershott, T., Menkveld, A. J., 2014. Price pressures. *Journal of Financial Economics* 114, 405–423.
- Huberman, G., Leshno, J., Moallemi, C., 2017. Monopoly without a monopolist: An economic analysis of the Bitcoin payment system. Unpublished working paper. Columbia Business School and University of Chicago.
- Jolliffe, I. T., 2002. *Principal Component Analysis*. Springer, New York.

- Kyle, A. S., 1985. Continuous Auctions and Insider Trading. *Econometrica* 53, 1315–1335.
- Nakamoto, S., 2008. Bitcoin: a peer-to-peer electronic cash system.
- Pagnotta, E., Buraschi, A., 2018. An equilibrium valuation of bitcoin and decentralized network assets. Unpublished working paper. Columbia Business School and University of Chicago.
- Raskin, M., Yermack, D., 2017. Digital currencies, decentralized ledgers, and the future of central banking. Unpublished working paper. NBER.
- Rosenthal, L., Young, C., 1990. The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever N.V./PLC. *Journal of Financial Economics* 26, 123–141.
- Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *Journal of Finance* 52, 35–55.

Appendix

Table A1.

Summary of exchanges.

This table summarizes some basic information of main exchanges in the paper. Margin trading includes both long and short positions. Trading fee can decrease with the size of trade, and the maker fee can be lower than taker fees.

Exchange	Headquarter	Supported fiat currencies	Trading rules	Allow margin trading?	Date allow margin trading	Trading fee rate	Time needed to withdraw/deposit fiat currencies
Binance	HK	USDT		No		0.02%-0.10%	
Bitfinex	HK	EUR, USD-USDT	If traders meet requirements of verification process, they can access all fiat currencies. (Note: Bitfinex currently does not serve US nationals.) Can transact in any fiat currency as a verified user but will need to clarify with the bank what currencies they will receive.	Yes	06/2013	0.00%-0.20%	1-10 business days
bitFlyer	Japan	JPY		Yes	11/2015	0.01%-0.15%	1-5 business days
Bithumb	Korea	KRW		No		0.08%-0.15%	1-3 business days
Bitstamp	UK	EUR, USD	Trading in EUR and USD is possible using the same account.	No		0.10%-0.25%	1-5 business days
Bittrex	USA	USDT		No		0.25%	
BtcBox	Japan	JPY		Yes	03/2014	0.05%	1-2 business days
BTCC	China	CNY		Yes	03/2014	0.20%	1-2 business days
BTC-e	Russia	EUR, RUR, USD	Trading in all supported fiat currencies is possible.	Yes	12/2015	0.20%	2-10 business days
Coinbase	USA	EUR, GBP, USD	Only support trading in the local fiat currency. Digital currency buys and sells are only available to residents of the US, Australia, Canada, and Europe.	No		1.49%	1-5 business days
Gemini	USA	USD		No		0.00%-0.25%	4-5 business days
Huobi	China	CNY		Yes	03/2014	0.20%	1-2 business days
Kraken	USA	CAD, EUR, GBP, JPY, USD	All trading pairs are accessible everywhere. People can trade on any currency pair regardless of the country of residence.	Yes	05/2015	0.00%-0.26%	1-5 business days
OKCoin	China	CNY		Yes	03/2014	0.20%	1-2 business days
Poloniex	USA	USDT		Yes	05/2015	0.00%-0.25%	
Quoine	Japan	AUD, EUR, IDR, JPY, SGD, USD	Only allow trading in base currency, which is the person's country of residence selected at registration. But people can change the base currency every 6 months.	Yes	09/2014	0.00%-0.25%	1-3 business days
Zaif	Japan	JPY		Yes	09/2017	-0.05%-0.01%	1 business day

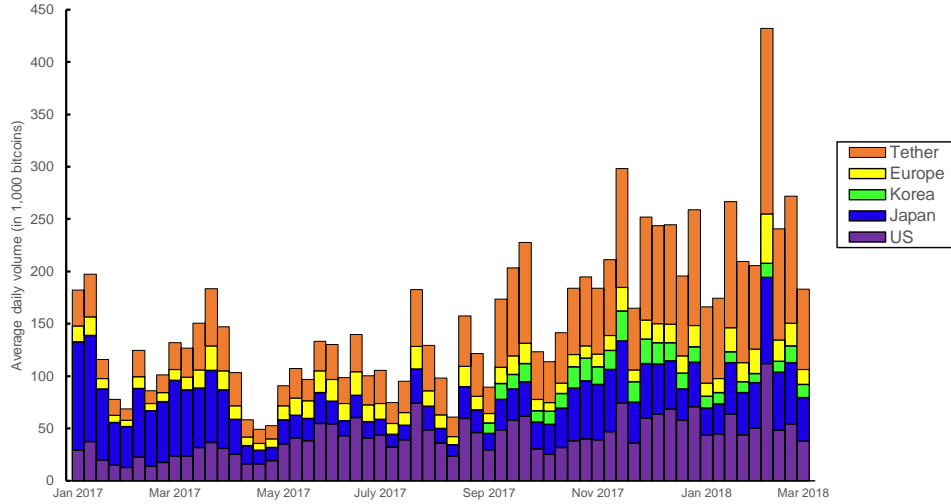
Table A2.

Correlation of arbitrage indexes across countries.

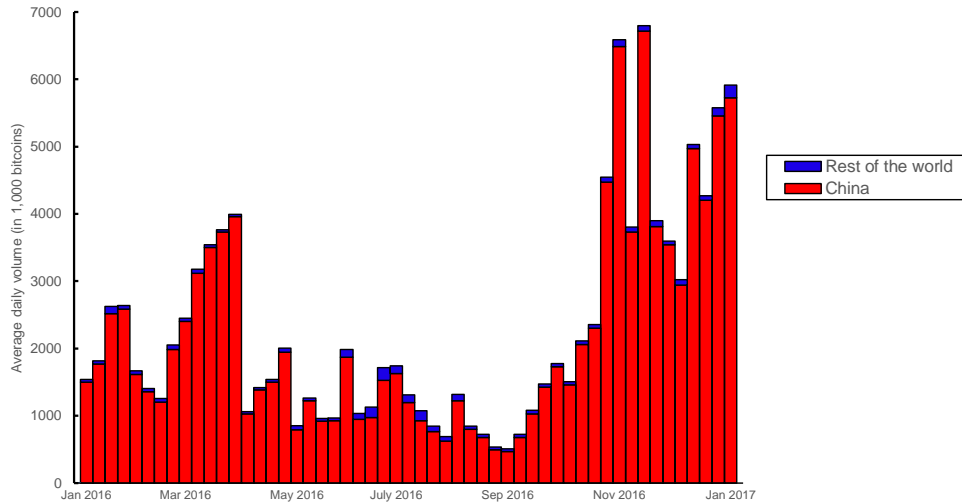
This table reports the correlation structure between all arbitrage indexes calculated at the minute level. The time period over which they are calculated is January 2017 to February 28, 2018. Fig. 6 in the paper is based on this correlation matrix.

	KRW	JPY	IDR	SGD	VND	AUD	PKR	ZAR	TRY	MXN	BRL	ILS	PLN	GBP	RUB	EUR	CAD
KRW	100																
JPY	78	100															
IDR	81	62	100														
SGD	76	81	71	100													
VND	81	71	79	79	100												
AUD	78	71	73	80	75	100											
PKR	57	35	74	57	65	54	100										
ZAR	52	44	63	54	45	56	57	100									
TRY	43	36	61	46	47	45	62	72	100								
MXN	65	67	52	51	47	49	39	49	42	100							
BRL	26	31	37	30	20	27	39	57	39	45	100						
ILS	36	24	49	45	42	43	67	41	40	33	37	100					
PLN	31	26	36	40	34	31	49	29	40	46	32	71	100				
GBP	19	3	35	18	21	23	45	27	24	-2	10	29	4	100			
RUB	29	13	32	35	52	33	54	2	15	6	-4	57	60	26	100		
EUR	7	6	10	19	13	17	18	2	0	5	2	45	53	23	56	100	
CAD	-17	-19	-2	-5	-8	-4	12	6	1	-8	20	38	35	15	40	58	100

Figures



Panel A



Panel B

Fig. 1. Total trading volume. This figure shows the average daily volume of bitcoin to fiat currency trading per week (reported in 1,000 BTC), from January 2017 until February 28, 2018 (Panel A) and from January 2016 to December 2016 (Panel B). The volume is reported across all 15 exchanges in the Kaiko data. We exclude any volume that is coin-to-coin trading. The fraction of volume that is generated on exchanges in different regions is indicated with different colors. The regions are China, Europe, Hong Kong, Japan, Korea, US, and Tether. In Panel A we report all regions apart from China. In Panel B we report all regions including China. The trading volume in China is indicated in red, while the volume in the rest of the world is in blue and was less than 5%.

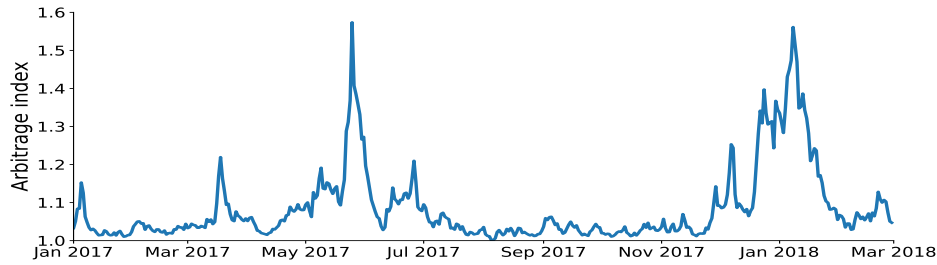
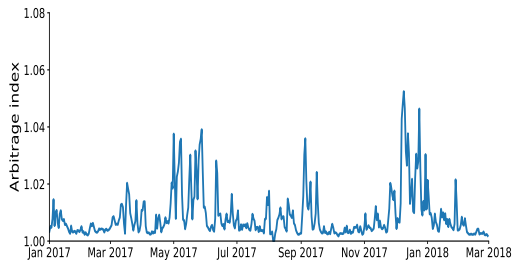
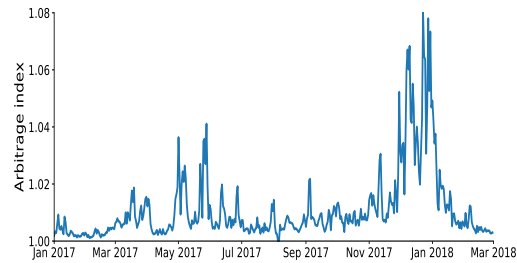


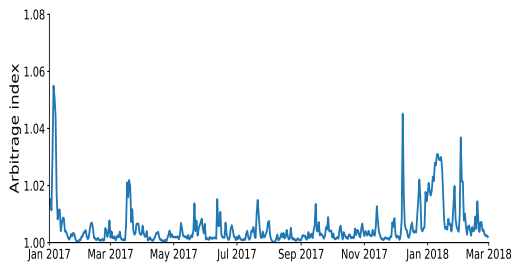
Fig. 2. Arbitrage index. The arbitrage index is calculated based on the volume-weighted price per minute for each exchange and averaged at the daily level. For a given minute the maximum volume-weighted price across all exchanges is divided by the minimum volume-weighted price in that minute. The set of exchanges include Binance, Bitfinex, bitFlyer, Bithumb, Bitstamp, Bittrex, Coinbase, Gemini, Kraken, Korbit, Poloniex, Quoine, and Zaif from January 2017 until February 28 2018.



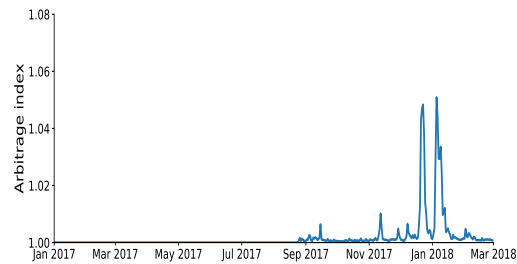
Panel A: USA



Panel B: Europe

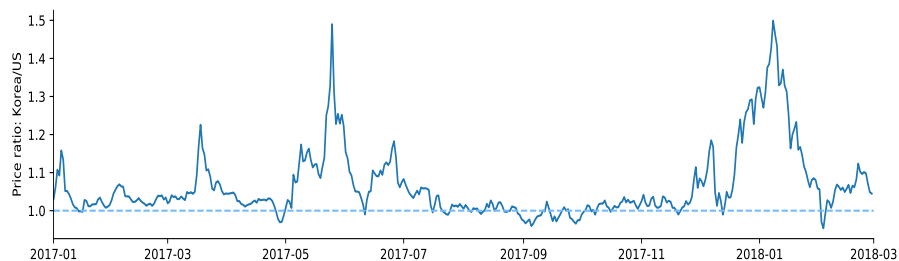


Panel C: Japan

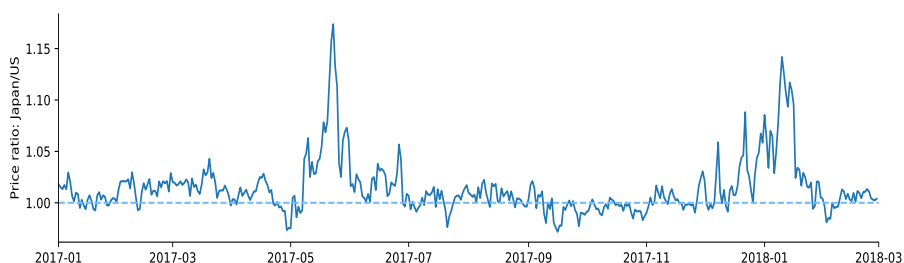


Panel D: Korea

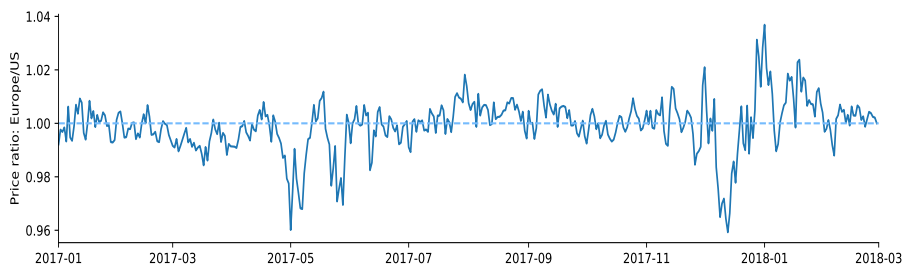
Fig. 3. Arbitrage index within regions. This figure calculates the arbitrage index for bitcoin to fiat currency on all the exchanges within a region from January 2017 until February 28, 2018. The arbitrage index is calculated based on the volume-weighted price of bitcoin per minute for each exchange and averaged at the daily level. For a given minute the maximum price across all exchanges is divided by the minimum price in that minute. Outliers are removed by replacing any price movement of more than 10% between two adjacent transactions. Panel A uses data from US exchanges: Bitstamp, Coinbase, Gemini, and Kraken: USD. Panel B uses data from European exchanges: Bitstamp, Kraken, and Coinbase: EUR. Panel C uses data from Japanese exchanges: bitFlyer, Quonie, and Zaif. Panel D uses data for Korean exchanges: Bithumb and Korbit.



Panel A: US vs. Korea

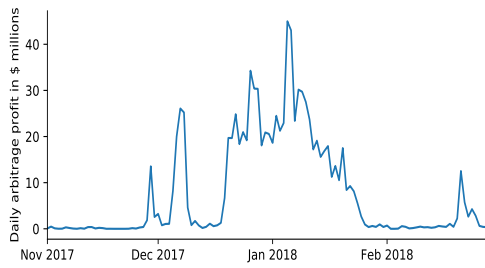


Panel B: US vs. Japan

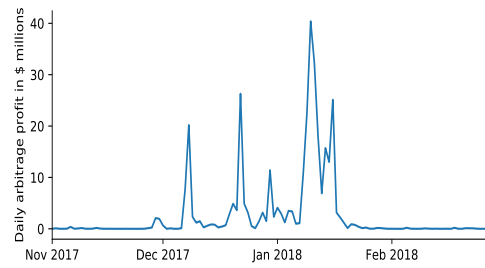


Panel C: US vs. Europe

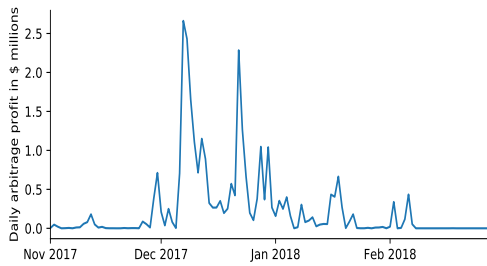
Fig. 4. Price ratio across regions. This figure plots the average price ratio between the price of bitcoin to USD across pairs of regions, from January 2017 until February 28, 2018. The ratio is calculated based on the volume-weighted price per minute across all the exchanges in a region and averaged at the daily level. Panel A plots the price ratio for the US versus Korea. Panel B repeats the same calculation for the exchanges in Japan versus the US. Finally, Panel C reports these calculations for the US versus Europe.



Panel A: US and Korea



Panel B: US and Japan



Panel C: US and Europe

Fig. 5. Daily profits. This figure plots the arbitrage profits between two regions calculated at the second level and then aggregated to the daily level. We include only price differences between exchanges if the price difference is larger than 2%. For each second, the aggregate amount of low priced volume that could have been sold in a high price region is subtracted from the sell-initiated volume in the region that has the highest price in a given second. Panel A shows the profit between the US and Korea, Panel B between the US and Japan, and, finally, Panel C between the US and Europe.

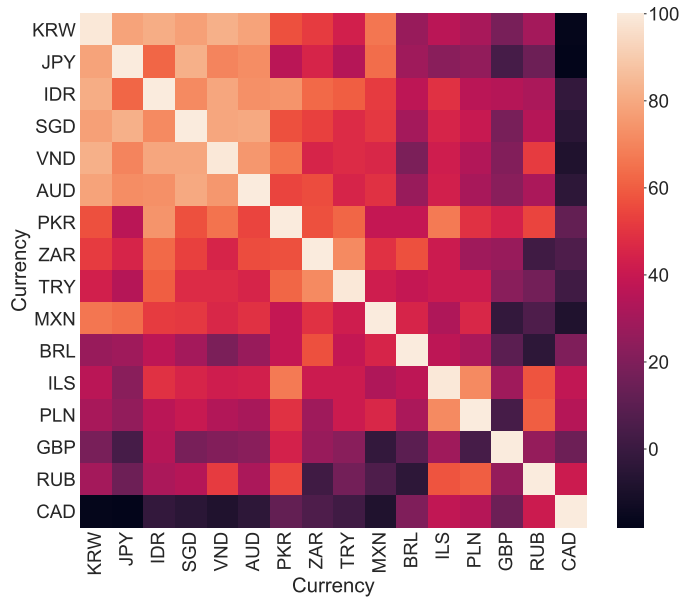


Fig. 6. Correlation matrix of arbitrage indexes across countries. For each of the countries in the data set we calculate the value weighted price of bitcoin across the exchanges in the given countries and form the arbitrage spread with respect to world market price at the second level. The table shows the correlation between the arbitrage index in the country on the left hand column to the country on the top of the matrix. The correlations are measured for the time period from January 2017 until the end of April 2018. The correlations are coded by a color scheme, where the lighter the color, the higher is the correlation. A legend is provided on the right side of the correlation matrix.

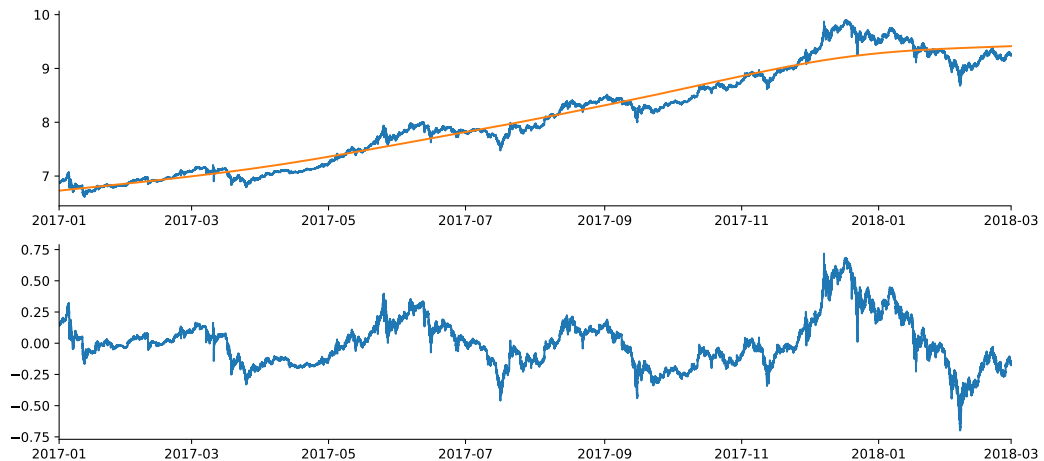


Fig. 7. Buying pressure. Panel A of this figure plots the series of the smoothed log bitcoin price in the US at the weekly level using the Hodrick-Prescott from the beginning of 2017 until April 2018. Panel B shows the deviations of the actual log bitcoin price from the smoothed log price.

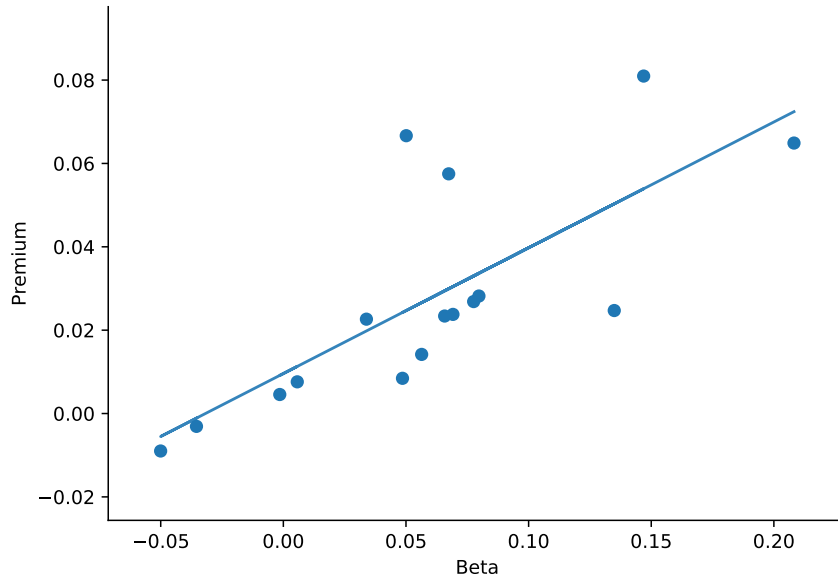


Fig. 8. Bitcoin premium. This figure shows the results from a cross-sectional regression of the volume-weighted average bitcoin price in a given country relative to the bitcoin price in the US (we call this the bitcoin premium since it is typically above the US bitcoin price) regressed on this country’s arbitrage beta. The variables are calculated from January 2017 to April 2018. We report the slope and the scatterplot for the countries in our data set. The estimated slope coefficient is 0.3 with a t -statistic of 4.6 and an adjusted R^2 of 0.56.

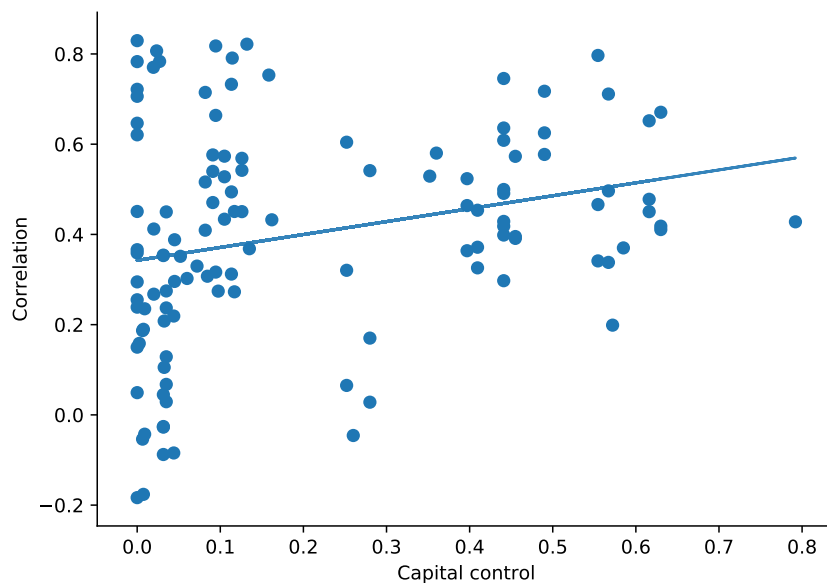
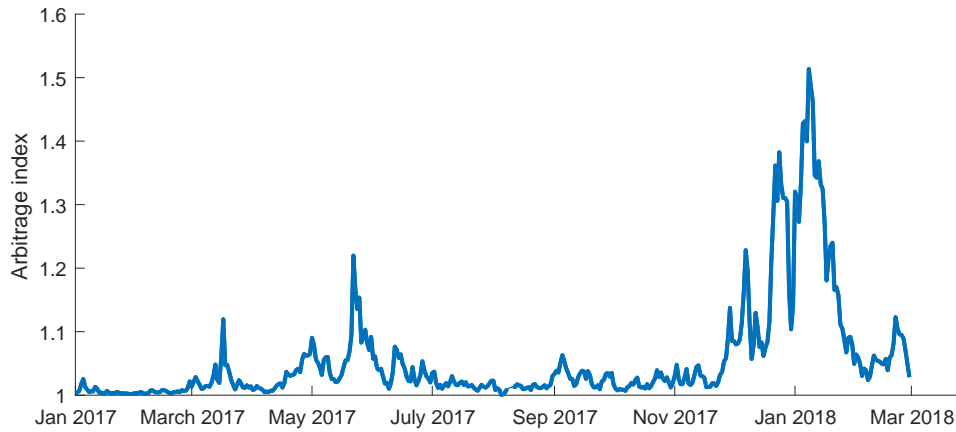
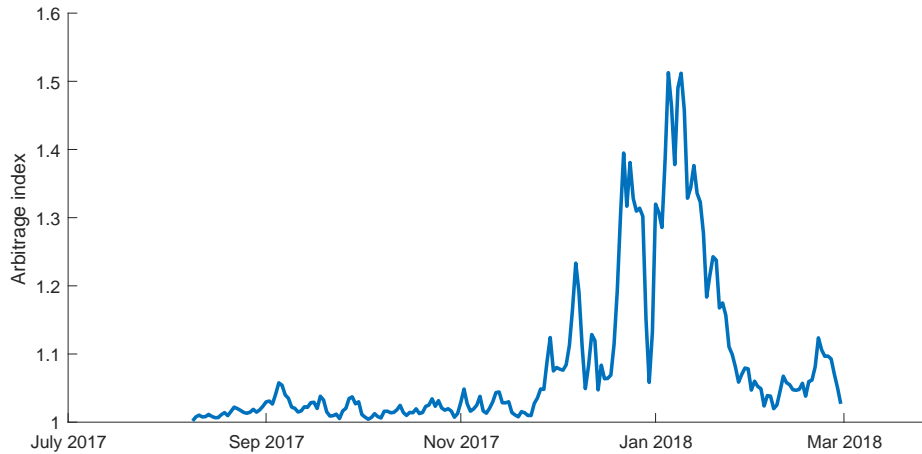


Fig. 9. Correlation in arbitrage spreads and capital control. The results in this figure are based on a data set of pairwise correlations in arbitrage spreads between two countries and a measure or pairwise capital controls, which is the product of the capital control index of the two countries. The capital control measure is based on the measure of capital market openness from the NBER database by [Fernandez et al. \(2015\)](#). The closer both countries are the closer the measure to one. The slope of this regression is 0.29 with a t -statistic of 3 and a R^2 of 0.07.

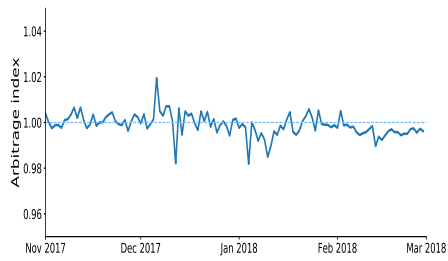


Panel A: Ethereum

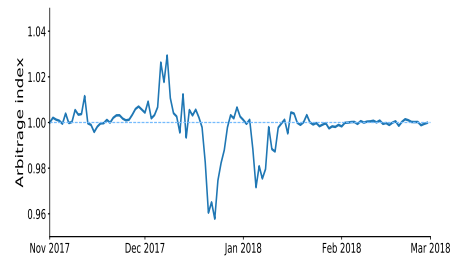


Panel B: Ripple

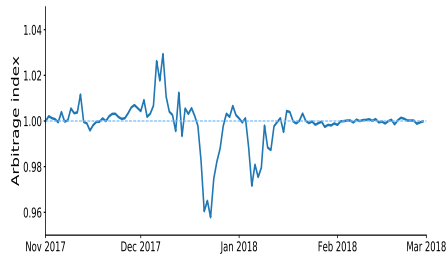
Fig. 10. Ethereum and ripple arbitrage index. This figure calculates the arbitrage index for ethereum and ripple to fiat currency on all the exchanges from January 2017 until February 28, 2018. Panel A shows the index for ethereum, Panel B for ripple. Each arbitrage index is calculated based on the volume-weighted minute price of the corresponding currency and is then averaged at the daily level. For a given minute the maximum price across all exchanges is divided by the minimum price in that minute.



Panel A: US vs. Japan



Panel B: US vs. Korea



Panel C: US vs. Europe

Fig. 11. Ethereum-bitcoin exchange rate across regions. This figure plots the average price ratio between the price of ethereum to bitcoin across pairs of regions from November 2017 until February 28th, 2018. The ratio is calculated based on the volume-weighted price per minute across all the exchanges in a region and averaged at the daily level. Panel A plots the daily price ratio for the US versus Japan. Panel B shows the results for the US versus Korea, and finally Panel C reports the results for US versus Europe.

Tables

Table 1.

Cryptocurrency exchanges.

This table reports the countries and names of exchanges that are included in the analysis and for which we have tick-level data. In the last column we report the source of the data. Data are provided either by Kaiko, a private data vendor, bitcoincharts.com, a public access website, or, in a few cases, we obtained data from the exchanges directly.

Currency	Country/Region	Exchange	Data source
AUD	Australia	Bitmarkets	Bitcoincharts.com
BRL	Brazil	Mercado	Exchange
		Foxbit	Bitcoincharts.com
CAD	Canada	Kraken	Kaiko
CNY	China	BTCC	Kaiko
		OkCoin	Kaiko
		Huobi	Kaiko
EUR	Euro	Bitstamp	Kaiko
		Coinbase	Kaiko
		Kraken	Kaiko
GBP	Great Britain	Coinbase	Kaiko
IDR	Indonesia	Btcoid	Bitcoincharts.com
ISL	Iceland	bit2cl	Bitcoincharts.com
JPY	Japan	Bitbox	Kaiko
		bitFlyer	Kaiko
		Coincheck	Bitcoincharts.com
		Fisco	Bitcoincharts.com
		Quoine	Kaiko
		Zaif	Kaiko
KRW	Korea	Bithumb	Kaiko
		Korbit	Bitcoincharts.com
MXN	Mexico	Bitso	Exchange
PKR	Pakistan	Urbit	Bitcoincharts.com
PLN	Poland	Bitbay	Bitcoincharts.com
		Bitmarket	Bitcoincharts.com
RUB	Russia	BTC-e	Kaiko
SGD	Singapore	Quoine	Kaiko
ZAR	South Africa	Bitx (Luno)	Bitcoincharts.com
TRY	Turkey	Koinim	Exchange
USD	USA	Bitstamp	Kaiko
		BTC-e	Kaiko
		Coinbase	Kaiko
		Gemini	Kaiko
		Kraken	Kaiko
VND	Vietnam	Vbtc	Bitcoincharts.com
Tether		Bitfinex	Kaiko
		Binance	Kaiko
		Bittrex	Kaiko
		Poloniex	Kaiko

Table 2.

Trading volume.

This table reports average daily trading volume, number of trades and number of trades for each exchange from tick data. We also add a last column in each panel where we report the average bid-ask spread based on data from the order book for each exchange. The first for columns cover the time period from January 2017 until July 31, 2017 and the second from August 2017 until February 28, 2018. Exchanges are organized by the five regions/currencies: US, Japan, Korea, Europe and Tether.

Exchange	Jan 2017—Jul 2017				Aug 2017—Feb 2018			
	Average daily volume (\$ millions)	Average daily number of trades (thousands)	Average size of trades (\$)	Average bid-ask spread (basis points)	Average daily volume (\$ millions)	Average daily number of trades (thousands)	Average size of trades (\$)	Average bid-ask spread (basis points)
<i>US</i>								
Coinbase: USD	19.51 (15.56)	27.86 (13.17)	619.45 (272.11)	3.04	180.47 (172.03)	83.90 (56.32)	1,862.88 (654.92)	0.94
Bitstamp: USD	19.92 (17.45)	11.69 (7.18)	1,496.07 (485.48)	12.34	132.11 (104.15)	41.49 (28.24)	3,022.68 (594.30)	12.81
Gemini	13.57 (15.77)	6.90 (5.88)	1,698.51 (778.51)	4.28	77.66 (68.20)	23.41 (14.54)	3,078.75 (988.45)	2.55
Kraken: USD	8.59 (7.69)	9.20 (7.09)	862.10 (209.92)	21.54	45.44 (33.12)	24.45 (14.08)	1,792.53 (616.32)	9.37
<i>Japan</i>								
bitFlyer	23.49 (11.71)	31.91 (16.53)	783.35 (267.74)	6.50	173.98 (128.90)	88.42 (59.72)	1,922.94 (503.69)	6.20
Zaif	6.73 (7.62)	78.42 (86.69)	98.41 (31.98)	3.42	77.30 (50.58)	163.27 (53.31)	472.76 (305.25)	3.78
Quoine	22.18 (19.57)	14.06 (8.26)	1,495.37 (563.06)	18.58	104.52 (113.65)	41.01 (36.15)	2,552.31 (1583.85)	11.03
<i>Korea</i>								
Bithumb	-	-	-	-	142.82 (81.69)	50.66 (16.76)	2,621.91 (814.16)	7.10
Korbit	10.73 (9.76)	6.00 (4.24)	1,567.06 (485.97)	-	49.06 (45.37)	15.85 (9.30)	2,761.54 (815.11)	-
<i>Europe</i>								
Kraken: Euro	19.14 (13.28)	20.71 (10.18)	856.02 (272.21)	8.88	78.20 (58.11)	38.71 (23.79)	1,788.59 (611.87)	9.23
Coinbase: Euro	2.05 (1.80)	8.65 (5.11)	213.64 (84.41)	10.48	33.99 (37.92)	42.95 (39.83)	675.67 (220.16)	3.06
Bitstamp: Euro	2.73 (2.72)	3.09 (2.34)	776.46 (290.93)	34.69	32.66 (31.98)	19.91 (17.04)	1,493.87 (490.92)	26.57
<i>Tether</i>								
Bitfinex	25.76 (16.70)	18.14 (8.62)	1,369.84 (375.17)	3.96	445.36 (330.12)	117.88 (70.81)	3,460.44 (871.99)	2.39
Poloniex	-	25.48 (17.08)	-	15.34	68.33 (55.30)	61.70 (43.29)	1,134.81 (442.97)	10.40
Binance	-	-	-	-	224.41 (186.30)	156.21 (146.21)	1,580.26 (590.98)	-
Bitrex	-	2.26 (4.67)	-	116.46	60.42 (47.34)	39.19 (23.97)	1,383.77 (368.23)	13.78

Table 3.

Summary statistics.

This table describes the higher moments of bitcoin returns at the daily, hourly, and five-minute level from January 2017 until February 28, 2018. For each frequency we report the annualized standard deviation, skewness and kurtosis of returns, as well as the autocorrelation at one, two and three lags and cross-correlation across exchanges. These statistics are calculated across all the exchanges in our data but without the Chinese exchanges due to data availability.

Return frequency	Std. dev	Skewness	Kurtosis	ρ_1	ρ_2	ρ_3	Cross corr
5-Minute	1.45	0	84.00	0.06	0	0	0.63
Hour	1.20	-0.11	15.82	0	-0.04	0	0.83
Daily	1.04	0.3	3.64	-0.02	0.01	0	0.95

Table 4.

Correlation of arbitrage indices.

This table reports the correlation structure between the regional arbitrage indexes in the US, Korea, Japan and Europe with the overall arbitrage index. These are the indexes calculated in Fig. 4 and 5. All arbitrage indexes are calculated at the minute. The time period over which they are calculated is January 2017 until February 28, 2018.

	US	Korea	Japan	Europe	All
US	100	15.3	15.3	59.6	37.5
Korea		100	29.8	31.7	54.4
Japan			100	18.0	42.3
Europe				100	49.8
All					100

Table 5.

Arbitrage betas.

This table reports the country-level arbitrage betas estimated over the time period from the beginning of 2017 until April 2018. Each column reports the beta for a given country. The arbitrage betas are estimated by regressing the country-level deviations of a country's bitcoin price relative to the US on a measure of buying pressure. The measure of buying pressure is obtained as the difference between the log bitcoin price in the US minus the smoothed log bitcoin price at the weekly level using the Hodrick-Prescott.

	KRW	JPY	IDR	SGD	VND	AUD	PKR	ZAR	TRY	MXN	BRL	ILS	PLN	GBP	RUB	CAD	EUR
Premium	0.066	0.015	0.025	0.008	0.024	0.028	0.058	0.081	0.024	0.027	0.067	0.023	0.005	0.008	-0.009	-0.003	-0.001
Beta	0.22	0.061	0.135	0.049	0.071	0.08	0.067	0.147	0.068	0.078	0.049	0.034	-0.002	0.006	-0.05	-0.036	-0.009

Table 6.

Factor analysis of signed volume.

This table reports the results of a factor analysis applied to signed volume data on 14 exchanges as described by model (1) in Section 7. Each panel reports the factor loadings, weights, and R -squared for the first common factor across our 14 main exchanges at the five-minute, hourly, and daily frequency, respectively.

	Bitfinex	Coinbase USD	Bitstamp USD	Gemini	Kraken USD	Kraken EUR	Coinbase EUR	Bitstamp EUR	bitFlyer	Quoine	Zaif	Bitfumb	Poloniex	Bittrex
5-min frequency														
Loadings	0.354	0.124	0.102	0.046	0.034	0.050	0.020	0.016	0.088	0.041	0.026	0.031	0.034	0.033
Weights	0.442	1.173	1.195	0.702	1.548	1.148	4.718	1.900	0.951	0.278	1.962	1.284	1.712	1.931
R^2	0.596	0.576	0.526	0.211	0.307	0.326	0.453	0.202	0.420	0.084	0.298	0.251	0.331	0.353
Hourly frequency														
Loadings	0.325	0.132	0.102	0.049	0.045	0.062	0.022	0.017	0.079	0.026	0.030	0.035	0.039	0.036
Weights	0.420	0.809	1.210	0.825	2.537	1.580	3.978	1.681	0.679	0.102	1.471	0.862	1.802	1.733
R^2	0.672	0.608	0.647	0.352	0.623	0.586	0.558	0.276	0.424	0.034	0.376	0.288	0.497	0.465
Daily frequency														
Loadings	0.312	0.125	0.106	0.052	0.052	0.071	0.015	0.019	0.073	0.016	0.032	0.042	0.042	0.042
Weights	0.375	0.326	1.268	1.488	3.256	1.698	1.791	1.669	0.370	0.052	1.709	0.519	2.205	1.992
R^2	0.669	0.393	0.704	0.559	0.758	0.677	0.294	0.330	0.297	0.013	0.471	0.256	0.608	0.580

Table 7.

Factor analysis of returns.

This table reports the result of factor analysis applied to log-return data from 14 exchanges as described by model (2) in Section 7. Each panel reports the factor loadings, weights, and R -squared for the first common factor across our 14 main exchanges at the five-minute, hourly, and daily frequency, respectively.

	Bitfinex	Coinbase USD	Bitstamp USD	Gemini	Kraken USD	Kraken EUR	Coinbase EUR	Bitstamp EUR	bitFlyer	Quoine	Zaif	Bitfumb	Poloniex	Bittrex
5-min frequency														
Loadings	1.121	1.023	1.033	1.036	0.697	0.697	0.929	0.972	0.840	0.926	0.828	0.819	1.077	1.062
Weights	0.162	0.112	0.121	0.164	0.032	0.030	0.048	0.050	0.051	0.024	0.021	0.032	0.099	0.054
R^2	0.885	0.818	0.832	0.876	0.442	0.425	0.615	0.639	0.608	0.436	0.378	0.486	0.804	0.676
Hourly frequency														
Loadings	1.033	0.986	1.003	1.005	0.964	0.963	0.966	0.990	0.892	0.949	0.905	0.853	1.040	1.077
Weights	0.137	0.117	0.142	0.153	0.062	0.041	0.031	0.077	0.022	0.023	0.017	0.015	0.098	0.064
R^2	0.962	0.953	0.962	0.966	0.908	0.865	0.825	0.928	0.754	0.773	0.708	0.664	0.946	0.919
Daily frequency														
Loadings	1.032	0.984	0.998	1.001	0.965	0.982	0.951	0.977	1.097	1.115	1.123	0.986	1.018	1.021
Weights	0.082	0.056	0.308	0.155	0.068	0.047	0.022	0.098	0.008	0.009	0.008	0.004	0.074	0.061
R^2	0.988	0.982	0.998	0.994	0.985	0.978	0.951	0.990	0.888	0.903	0.887	0.796	0.987	0.984

Table 8.
Price impact.

This table reports the results from time-series regressions of the common component of returns on the contemporaneous and lagged common component in signed volume extracted using data from our 14 main exchanges:

$$r_t^* = \lambda s_t^* + \sum_{\tau=1}^T \lambda_{\tau} s_{t-\tau}^* + \varepsilon_t.$$

The first three columns report the estimates at the 5-minute frequency; the next six columns show results at the hourly and then daily frequencies. T -statistics are computed using the MacKinnon and White's (1985) heteroskedasticity robust standard errors and are given in parentheses.

	5-min frequency			Hourly frequency			Daily frequency		
	$\lambda \times 10^4(\%)$			$\lambda \times 10^4(\%)$			$\lambda \times 10^4(\%)$		
s_t^*	8.8	9.9	10.1	6.0	6.6	6.6	3.6	3.9	4.0
	(80.06)	(86.19)	(88.05)	(35.12)	(39.7)	(40.41)	(16.92)	(19.93)	(18.96)
s_{t-1}^*		-3.1	-2.6		-2.1	-2.0		-1.1	-1.1
		(-36.54)	(-32.24)		(-16.53)	(-15.67)		(-4.05)	(-3.62)
s_{t-2}^*			-0.8			-0.4			-0.0
			(-11.68)			(-3.71)			(-0.2)
s_{t-3}^*			-0.5			-0.1			-0.1
			(-7.56)			(-1.22)			(-0.76)
s_{t-4}^*			-0.4			-0.3			-0.3
			(-6.88)			(-3.00)			(-1.71)
s_{t-5}^*			-0.3			-0.1			0.3
			(-5.24)			(-1.33)			(1.57)
R^2	0.54	0.60	0.61	0.6	0.66	0.67	0.69	0.75	0.76
N	46666	46666	46666	4085	4085	4085	179	179	179

Table 9.

Idiosyncratic signed volume.

This table reports the results from time-series regressions of the idiosyncratic component of the signed volume on each of the exchange list on the top of the column, regressed on the deviation of the price from the common price component and past three lags of the idiosyncratic component of the signed volume of the same exchange. The idiosyncratic components, \hat{s}_{it} and \hat{p}_{it} , are obtained as the residual values of signed volume and prices after taking out the common component from each.

$$\hat{s}_{it} = \gamma_i \hat{p}_{it-1} + b_{1i} \hat{s}_{it-1} + b_{2i} \hat{s}_{it-2} + b_{3i} \hat{s}_{it-3} + \varepsilon_{it}.$$

T-statistics are computed using the MacKinnon and White's (1985) heteroskedasticity robust standard errors and are given in parentheses.

	Bitfinex	Coinbase USD	Bitstamp USD	Gemini	Kraken USD	Kraken EUR	Coinbase EUR	Bitstamp EUR	bitFlyer	Quoine	Zaif	Bithumb	Poloniex	Bittrex
	5-min frequency													
$\gamma_i \times 10^{-2}$	-1.12 (-2.4)	0.86 (4.12)	-1.29 (-4.9)	-0.45 (-2.66)	-1.61 (-15.41)	-0.42 (-4.54)	0.04 (1.27)	-0.22 (-5.16)	-0.49 (-7.78)	0.27 (2.92)	-0.15 (-7.23)	0.05 (7.12)	0.14 (3.18)	-0.23 (-5.35)
b_{1i}	0.09 (5.57)	0.17 (17.72)	0.08 (6.1)	0.06 (6.18)	0.07 (6.58)	0.09 (8.39)	0.13 (11.6)	0.15 (11.73)	0.15 (17.55)	0.22 (16.83)	0.23 (21.31)	0.21 (32.29)	0.13 (9.9)	0.15 (13.42)
b_{2i}	0.06 (6.52)	0.08 (9.24)	0.07 (6.15)	0.02 (2.85)	0.03 (2.74)	0.05 (4.64)	0.05 (5.35)	0.04 (3.91)	0.05 (6.76)	0.08 (7.52)	0.07 (8)	0.09 (14.23)	0.06 (4.67)	0.04 (4.71)
b_{3i}	0.05 (5.19)	0.08 (9.82)	0.04 (5.13)	0.02 (3.04)	0.02 (2.32)	0.03 (3.7)	0.08 (8.8)	0.04 (5.34)	0.06 (7.9)	0.09 (8.18)	0.07 (7.87)	0.08 (12.34)	0.03 (2.52)	0.06 (7.01)
R^2	0.02	0.05	0.01	0	0.01	0.01	0.03	0.03	0.04	0.09	0.08	0.08	0.02	0.03

Table 10.

Idiosyncratic returns.

This table reports the results from time-series regressions of the deviation of the price from the common price component on each of the exchange list on the top of the column, regressed on past three lags of the deviation of the price from the common price component and the idiosyncratic component of the signed volume of the same exchange. The idiosyncratic components, \hat{s}_{it} and \hat{p}_{it} , are obtained as the residual values of signed volume and prices after taking out the common component from each.

$$\hat{p}_{it} = \lambda_i \hat{s}_{it} + a_{1i} \hat{p}_{it-1} + a_{2i} \hat{p}_{it-2} + a_{3i} \hat{p}_{it-3} + \varepsilon_{it}.$$

T-statistics are computed using the MacKinnon and White's (1985) heteroskedasticity robust standard errors and are given in parentheses.

	Bitfinex	Coinbase USD	Bitstamp USD	Gemini	Kraken USD	Kraken EUR	Coinbase EUR	Bitstamp EUR	bitFlyer	Quoine	Zaif	Bitfumb	Poloniex	Bittrex
	5-min frequency													
$\lambda_i \times 10^4(\%)$	2.86 (16.49)	17.35 (22.83)	5.76 (9.18)	8.37 (14.35)	40.95 (21.14)	41.66 (27.66)	172.03 (25.64)	15.8 (7.43)	17.13 (22.26)	4.35 (6.58)	59.61 (13.34)	32.1 (25.13)	20.1 (12.28)	22.66 (14.00)
a_{1i}	0.6 (48.44)	0.63 (16.28)	0.55 (56.57)	0.59 (34.58)	0.56 (43.48)	0.63 (40.07)	0.73 (29.02)	0.5 (25.25)	0.83 (40.69)	0.79 (26.36)	0.84 (14.73)	0.83 (50.95)	0.61 (54.99)	0.6 (61.34)
a_{2i}	0.23 (17.07)	0.18 (5.58)	0.23 (21.47)	0.24 (13.5)	0.2 (14.75)	0.19 (11.51)	0.16 (4.18)	0.26 (18.78)	0.12 (4.8)	0.15 (5.55)	0.01 (0.08)	0.12 (6.45)	0.21 (19.32)	0.21 (21.32)
a_{3i}	0.16 (12.84)	0.18 (5.51)	0.2 (21.89)	0.16 (11.18)	0.21 (19.68)	0.16 (9.62)	0.1 (4.1)	0.23 (13.54)	0.04 (1.79)	0.05 (2.59)	0.15 (3.3)	0.05 (3.64)	0.17 (16.56)	0.18 (18.78)
R^2	0.98	0.97	0.94	0.96	0.89	0.95	0.98	0.95	0.99	0.98	0.98	0.99	0.99	0.98