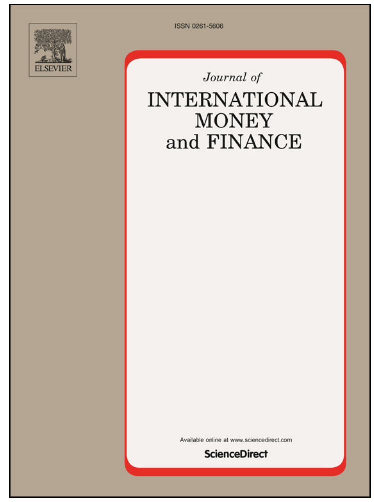
Journal Pre-proof

Cryptomarket Discounts

Nicola Borri and Kirill Shakhnov



PII: S0261-5606(23)00164-X
DOI: https://doi.org/10.1016/j.jimonfin.2023.102963
Reference: JIMF 102963

To appear in: *Journal of International Money and Finance*

Please cite this article as: N. Borri and K. Shakhnov, Cryptomarket Discounts, *Journal of International Money and Finance*, 102963, doi: https://doi.org/10.1016/j.jimonfin.2023.102963.

Highlights

- Large, persistent bitcoin price differences: Bitcoin prices diverge significantly across exchanges and currency pairs, posing questions about market efficiency and cross-border payment systems.
- Comprehensive analysis of bitcoin price dispersion: The paper considers 135 global exchanges and documents the distribution of daily bitcoin prices, revealing significant and varying discounts for both fiat and crypto pairs.
- Importance of location: location component accounts for at least 50 percent of this total variability for fiat pairs.
- Market segmentation: stricter capital controls increase discount variability, amplifying local supply-demand shocks, proxied by mining activities and investor attention.

Cryptomarket Discounts *

Nicola Borri Kirill Shakhnov

LUISS Surrey

This version: September 18, 2023

Abstract

This paper studies the efficiency of the cryptocurrency market by looking at the distribution of bitcoin prices over time and across exchange-currency pairs. We document persistent differences in relative bitcoin prices (or discounts), with a half-life of 1 day, and a distribution which is leptokurtic, skewed to the right, with a standard deviation of 3.9%. The variability of discounts is larger in countries with tighter capital controls due to the combined effect of market segmentation and local supply and demand shocks, which we relate to location-specific mining activities and investor attention.

Keywords: cryptocurrency; limits to arbitrage; mining; multi-market trading

JEL Classification: G14, G15, F31

*Borri: Department of Economics and Finance, Luiss University, Viale Romania 32, 00197 Rome, Italy; nborri@luiss.it; Tel: +39 06 85225959; www.nicolaborri.com. Shakhnov: School of Economics, University of Surrey, Guildford, Surrey, GU2 7XH, UK; k.shakhnov@surrey.ac.uk; Tel: +44 (0)1483 689380; sites.google.com/site/kshakhnov.
We thank Bob Chirinko, Juan Passadore, Anton Tsoy, Sergei Kovbasyuk, Federico Carlini, Paolo Vitale, and seminar participants at Luiss, EIEF, Crypto Valley Conference on Blockchain Technology, Technische Universität Berlin, Bank of Italy. Aleksandr Shneider provided excellent research assistance. A separate Online Appendix is available on the authors' websites.

1 Introduction

Investors buy bitcoin on a multitude of exchanges, located in different countries, and against different fiat and cryptocurrencies. The notable large and persistent differences in bitcoin prices across these exchange and currency pairs question the efficiency of both cryptocurrency markets and cross-border payment systems and could hamper the growth of the derivatives cryptocurrency market and of cryptocurrency ETFs, which require reliable underlying price indices.

In this paper, we consider 135 exchanges around the globe, where investors can trade bitcoin for different fiat and cryptocurrencies (henceforth also fiat and crypto pairs, respectively), and establish the shape of the distribution of daily bitcoin prices over time, and across different exchanges and currencies. While the typical price distribution is roughly symmetric for bitcoin-to-crypto pairs, and more positively skewed for bitcoin-to-fiat pairs, for all pairs is leptokurtic, with a mean standard deviation of approximately 3.9% percent. The existence of exchange location-specific limits-to-arbitrage is a necessary, but not sufficient, condition for the existence of bitcoin price differences. We document that bitcoin price differences are larger and more persistent for fiat pairs and in exchanges located in countries with more severe capital controls, which limit arbitrage activity. For these pairs, we establish that local demand and supply shocks account for a large fraction of their time-series and cross-sectional variation in discounts. Describing the properties of these shocks and establishing their relation to observable factors is the main contribution of this paper.

Cryptocurrency is a novel asset class, with a total market capitalization of 1.2 trillion U.S. dollars as of July 2023 (it reached a peak of around 3 trillion U.S. dollars at the end of 2021). We focus on bitcoin because it was the first cryptocurrency, created in 2009 using a scheme proposed by Nakamoto (2008), and it currently accounts for one-half of both the total market capitalization and the trading volume according to data from CoinMarketCap. Bitcoin started trading in 2010 on the now-defunct Mt. Gox exchange. Presently, it is traded across numerous exchanges worldwide, operating on a 24-hour, seven days a week basis. A considerable number of these exchanges came into existence in 2014, which also serves as the starting point for our data collection.

Since bitcoin is a fungible asset, at a given point in time, and in competitive markets with no limits-to-arbitrage (e.g., transaction costs or capital flow restrictions), its price expressed in the same currency should be equal across different exchanges or currency pairs. In fact, the dollar prices of one bitcoin differ substantially. We refer to these differences, with respect to a baseline value, as *discounts*. Bitcoin is not the only asset to trade at a discount in different markets. Researchers and investors have looked at the price differences for cross-listed and home-market shares to study limits-to-arbitrage and market efficiency. For example, for Chinese companies that issue both A-shares in mainland China, restricted to mainland Chinese investors, and H-shares in Hong-Kong, available to international investors, Wang and Jiang (2004) document a large and time-varying discount, on average equal to almost 70 percent. In the case of American Depositary Receipts (ADRs) in U.S. markets, Gagnon and Karolyi (2010) find smaller but volatile average discounts of about 4.9 basis points that can reach large extremes. However, these examples refer to assets traded on a relatively small number of markets, or by a relatively small number of investors. In contrast, bitcoin is a truly global and fungible asset, traded in a multitude of exchanges, and by a large number of investors. One of the conditions for the cryptocurrency market for derivatives and De-Fi to keep growing, and for cryptocurrency ETFs to efficiently operate, is the availability of a single and reliable bitcoin price index that could serve as underlying security for a multitude of contracts or as a benchmark. The existence of bitcoin price differences across exchanges and currency pairs undermines the reliability of such an index.[1]

First, we document large, time-varying, and persistent price differences in bitcoin prices across exchanges and currencies using a sample substantially larger than the one considered in Makarov and Schoar (2020). Specifically, we consider 39 bitcoin-to-fiat pairs and 9 bitcoin-to-crypto pairs

[1]For this reason, exchanges that offer derivative contracts on cryptocurrency currently specify which quotes (e.g., the quotes for a small set of exchanges) they use to compute the price index underlying these contracts and what they would do in the event these quotes diverge by a significant amount. See, for example, https://www.binance.com/en/support/faq/547ba48141474ab3bddc5d7898f97928. While the practice of computing the price index using an average of different quotes has not prevented the derivatives and De-Fi markets to grow fast in the recent years, we cannot observe the growth of these markets, and the participation by institutional investors, under a counterfactual scenario in which bitcoin discounts were negligible. See, for example, Shiller (2008) on the importance of a reliable index for a well-functioning derivatives market.

traded on 135 exchanges located in 42 different countries. Although bitcoin-to-fiat pairs have a larger price dispersion, with an average standard deviation of bitcoin discounts of 5.3 percent per day, also for bitcoin-to-crypto pairs we observe daily discounts as large as 30 percent in absolute value.[2] In fact, the price distribution for all pairs follows a leptokurtic pattern with heavy tails. We decompose the cross-sectional variability of discounts into four components listed by order of importance: time, location, quality and currency. The cross-sectional standard deviation of bitcoin discounts exhibits significant variation over time, spanning from 0.1 to 12 percent for fiat pairs. The location-specific component, captured by the different exchange locations, explains, on average, more than 50 percent of this total variability for fiat pairs. While the relative contribution of the location-specific component has been subject to fluctuations over time, our findings indicate that it tends to be more significant when the overall variance of bitcoin discounts is higher. Further, the location-specific component is particularly important to explain the variance of the discounts of the fiat pairs while, for the crypto pairs, the currency component is the largest contributor to the overall variance. The latter captures the dispersion of bitcoin prices across currency-pairs *within* each location. Overall, the documentation of the properties of bitcoin discounts and the establishment of their decomposition into different components represent a novel contribution of this paper.

Second, in order to establish the importance of various contributing factors to the overall explanation of bitcoin price dispersion, we collect and merge data from multiple sources to provide a novel and comprehensive summary of stylized facts for cryptocurrency markets. We consider both traditional factors, like liquidity, and crypto-specific factors, like counter-party risks or blockchain factors. Since the focus of our paper is on the dispersion of bitcoin prices across exchanges and currencies, most of these measures are location-specific. We capture liquidity with data on trading volume and bid/ask spreads for the different exchange-currency pairs. Because most of cryptocurrency trading is *off*-the-blockchain and occurs in centralized cryptocurrency exchanges, investors are exposed to the risk that one party (i.e., the exchange) defaults on the transaction, like in the

[2] Bitcoin prices for some exchange-currency pairs can also be above the baseline value. In what follows a "positive" discount for a given pair, measures a *premium*.

recent scandal linked to the bankruptcy of FTX. We measure counter-party risks using data on exchange hacks and bankruptcies, regulatory shocks, and exchange wallets. The latter are, in some sense, like deposits for commercial banks. In our sample, on average, 10% of the pairs are not available for trade on any given day because of exchange hacks or software malfunction, and a total of 2.4 billion U.S. dollars worth of cryptocurrency was stolen. Finally, we consider several blockchain and cryptocurrency factors to capture spatial differences. For example, while mining activity was mostly concentrated in China, and now in the U.S., with miners collecting approximately 10 million U.S. dollars per day, trading is mostly concentrated in Japan and South Korea, which, together and since January 2018, account for 43% of the total. These disparities are likely to translate into a dispersion of bitcoin prices in the presence of partially segmented markets.

Third, in order to investigate the determinants of location-specific bitcoin discounts, we estimate panel regressions and factors at the exchange-location level. Local demand and supply shocks could contribute to differences in bitcoin prices in the presence of limits-to-arbitrage. In the data, we consider separately exchanges in "closed" and "open" economies under the assumption that limits-to-arbitrage are more likely to be severe in the former. Following Makarov and Schoar (2020), we classify economies into closed and open according to the K-control index of tightness of capital controls constructed by Fernández et al. (2016). The results of the panel estimates confirm that bitcoin discounts in closed economies are larger, more persistent and more sensitive to local demand and supply shocks. Specifically, shocks to local mining activity, our proxy for supply, are associated with lower relative bitcoin prices, while shocks to local Google searches, our proxy for demand, are associated with higher relative bitcoin prices. Further, we find that the effects of local demand and supply shocks in closed economies persist up to three weeks.

We build a simple model to guide the interpretation of our empirical results. The model is based on three assumptions. First, investors have heterogeneous beliefs about bitcoin values, like in Cong et al. (2021b) and Biais et al. (2023). This assumption gives investors a reason to trade. Second, markets are segmented, at least in the short run. This assumption generates price dispersion across markets and is motivated by the results in Makarov and Schoar (2020). Third, a market-maker infrequently and randomly visits each market and eliminates the price differences across these markets. This assumption guarantees that markets are integrated in the long run. In this framework, in the short run, local discounts are driven by changes in local demand and supply, while in the long run, prices are equalized across markets. We calibrate our model using mining activity, as a proxy for local supply, and Google searches, as a proxy for local demand, and show that it replicates the observed dispersion in bitcoin prices and the higher persistence of bitcoin discounts in locations with more severe limits-to-arbitrage. Our framework is related to existing models of bitcoin valuation. Specifically, like Pagnotta and Buraschi (2018), we emphasize the role of miners; and like Biais et al. (2019, 2023), we obtain that prices can be very volatile and unrelated to fundamentals. Differently from existing work, we explicitly model bitcoin price differences across markets and emphasize the role of local, as opposed to global, demand and supply.

This paper contributes to two strands of the literature. The first is the recent and growing empirical literature on cryptocurrency.[3] Part of this literature has documented the risk-return characteristics of cryptocurrency, typically considering returns before transaction costs. Liu and Tsyvinski (2021) find that only crypto-specific risk factors, like attention and bitcoin momentum, can contribute to account for the time-series risk-return relation in cryptocurrency. Liu et al. (2022) show that the cross-section of returns in cryptocurrency can be accounted for by three crypto-specific factors: the cryptocurrency market, size, and momentum. Shams (2020) documents a persistent structure in cryptocurrency returns, which is mostly explained by similarity in the investor bases, proxied by their trading location. The second part of this literature has studied the efficiency and pricing of cryptocurrency. Makarov and Schoar (2020), using a sample of pairs with small bid/ask spreads, document large arbitrage opportunities across exchanges in different locations and attribute them to market segmentation due to capital controls and weak financial

[3] Yermack (2013), Velde et al. (2013), and Dwyer (2015) are excellent primers that describe the functioning of the blockchain and cryptocurrency. Catalini and Gans (2016), Cong et al. (2021a), Ma et al. (2018), and Chiu and Koeppl (2019) analyze, from the perspective of economic theory, how blockchain technology and cryptocurrency will influence the rate and direction of innovation, the incentives and equilibria behind the "proof-of-work" protocols, and the settlement of securities.

institutions. Krückeberg and Scholz (2020) attribute these price differences to market inefficiencies and untapped arbitrage opportunities. In contrast, Borri and Shakhnov (2022), consider fiat and crypto pairs, traded in reliable exchanges located in countries with a low level of capital controls, account for all the transaction costs, and find that investors betting on the persistence or mean-reversion of these price differences are exposed to systematic risk.[4] This paper considers a larger sample, including the most traded fiat and cryptocurrencies, and emphasizes the importance of all contributing factors. While market segmentation explains the persistence of the largest price differences, local demand and supply shocks contribute to the variability of relative bitcoin prices. The second strand is the large finance literature on market efficiency and anomalies as well as limits to arbitrage. Some papers argue that market frictions can explain differences in the prices of homogeneous assets; others attribute them to differences in risk. Examples of the former, are Lee et al. (1991); Chen et al. (1993) for closed-end funds; Lamont and Thaler (2003) for tech stock carve-outs; Froot and Dabora (1999); Wang and Jiang (2004) for "Siamese twins"; Gagnon and Karolyi (2010) for cross-listed stocks, such as ADRs; Burnside (2011) for the forward premium; Du et al. (2018) for the deviations from the covered interest parity. Examples of the latter are Cochrane (2002) for tech stock carve-outs; Krishnamurthy (2002) for on-the-run and off-the-run bonds; Lustig and Verdelhan (2007) for the forward premium; De Jong et al. (2009) for dual-listed stocks. This paper focuses on the dispersion of bitcoin prices, for a large sample of fiat and crypto pairs, across markets, currencies, and over time, and relates price differences to frictions in cryptocurrency markets.

The rest of the paper is organized as follows: Section 2 describes the data; Section 3 introduces bitcoin discounts, our measure of price dispersion, and presents a decomposition in components related to time, exchange location, and currency pairs; Section 4 analyzes the determinants of these components using time-series and panel regressions; Section 5 presents a model with hetero-

[4]Brandvold et al. (2015) is an early paper which studies bitcoin price differences across exchanges using a short pre-2014 sample. The objective of the paper is to understand in which exchange price discovery takes place and identifies the latter with the now defunct Mt.Gox. Makarov and Schoar (2019) studies bitcoin price discovery across exchanges located in different countries using an early sample and highlights the importance of frictions *across* countries rather than *within* the same country. Huang et al. (2022) argue that triangular arbitrage opportunities between fiat and crypto pairs are higher at times of higher volatility in the equity market.

geneous investors, partially segmented markets, and a slow-moving market-maker, to interpret the empirical results. Finally, Section 6 presents our conclusions.

2 Bitcoin Data

Investors can purchase bitcoin using fiat or cryptocurrencies in different exchanges across the globe. There are mainly two types of exchanges.[5] The first type contains exchanges on which only crypto pairs are traded (e.g., bitcoin for ethereum), and where investors can deposit and withdraw only cryptocurrency; the second type, instead, contains exchanges where it is possible to trade fiat for cryptocurrencies (e.g., U.S. dollar for bitcoin), and where investors can deposit and withdraw both fiat and crypto. Crypto exchanges operate every day 24/7, including Saturdays, Sundays, and holidays, and use the UNIX time-stamp to track time and ensure immediate comparability of market prices. We collect bitcoin price and volume data on all exchanges listed on the data aggregator Cryptocompare. The longest sample is for the period January 1, 2014, to April 18, 2023. However, the samples differ for different pairs, which might enter or exit the sample, and tend to be shorter for crypto pairs, while the overall number of exchanges and pairs increases over time.[6] Table 1 reports information about the initial raw data (see Panel A). We start off a large sample containing data for 394 bitcoin-to-fiat pairs and 411 bitcoin-to-crypto pairs traded, respectively, on 147 and 127 exchanges. The number of pairs is larger than the effective number of fiat and cryptocurrencies, as each pair can be traded on more than one exchange. For example, we treat differently the bitcoin-to-dollar pair traded on the Kraken exchange and the bitcoin-to-dollar pair traded on the Bitfinex exchange. Similarly, we treat differently the bitcoin-to-dollar and the

[5]Recently, a further type of exchanges is gaining importance. It is the so called decentralized exchange, that allows peer-to-peer transactions without the need of an intermediary. Among the most notable decentralized exchanges are Uniswap and dYdX. Decentralized exchanges account for a relatively small share of total cryptocurrency trading and an even smaller share throughout our sample. For details on these exchanges see Barbon and Ranaldo (2021) and Lehar and Parlour (2021). In our paper, we exclude decentralized exchanges and focus, instead, on the so called centralized exchanges which we refer, simply, as cryptocurrency exchanges.

[6]We compute end-of-day prices and daily volume of transactions corresponding to 16:00 GMT, and drop observations corresponding to Saturdays, Sundays, and additional non-business days, to match daily bitcoin prices in all markets to daily spot rates for fiat currencies from WM/Reuters corresponding to 16:00 GMT. For cryptocurrencies, the exchange rate is the volume-weighted dollars per unit of cryptocurrency. We collect all the pairs with the U.S. dollar for each cryptocurrency and construct the exchange rate as volume-weighted relative price.

bitcoin-to-euro pairs both traded on the Kraken exchange. In fact, the original raw data contain only 45 fiat currencies and 14 cryptocurrencies. Note that the majority of coin were created during the ICO boom of 2018, but we focus on historically important coins with long trading histories.

Table 1: Our sample

	Currencies	Exchanges	Pairs	Observations
Panel A: Raw original data				
Fiat	45	147	394	418911
Crypto	14	127	411	555785
Panel B: Final sample after data cleaning				
	Currencies	Exchanges	Pairs	Observations
Fiat	39	135	299	348529
Crypto	9	61	224	340681

Notes: This table reports information on the total number of currencies, exchanges, exchange-currency pairs, and daily observations for the original raw data (panel A) and the final sample, after data cleaning (panel B). In each panel, the rows labeled "Fiat" refer to bitcoin-to-fiat currency pairs and those labeled "Crypto" to bitcoin-to-crypto currency pairs. Details on the data cleaning and the source of the data are reported in Section 2. The sample period is 1/1/2014–4/18/2023.

We restrict our sample along several dimensions to avoid the risk that data from less reliable exchanges or less liquid pairs could drive our empirical results. First, we eliminate exchanges where only crypto pairs are traded, because the reputability of these exchanges, and the quality of the information they provide to investors, has been questioned (for example, see Bitwise, 2019). Second, we eliminate all observations corresponding to days in which the trading volume, for a given pair, is equal to zero or missing, as these occur in correspondence to temporary shut-downs of the exchanges, for example, because of a cyberattack, a software maintenance or malfunction. Third, we exclude currency pairs with less than 31 observations; observations corresponding to a mean trading volume, in the previous week, smaller than 0.1 bitcoin; and the first available observation for each pair. Fourth, in order to avoid the possible influence of a small number of outliers on measured bitcoin price dispersion, we exclude observations corresponding to daily changes between the bitcoin price in any given pair and the volume-weighted average price being larger, in absolute value, than 50 percent. Finally, we drop bitcoin-to-gold pairs, as the latter is a commodity, and data from LocalBitcoins and other peer-to-peer platforms. Panel B of

Section 1 reports the total number of currencies, exchanges, pairs, and observations, for our final sample after data cleaning. This sample contains data for 39 fiat and 9 cryptocurrencies traded, respectively, on 135 and 61 exchanges. The number of sample pairs is larger than the number of fiat and cryptocurrencies. Specifically, the final sample contains 299 bitcoin-to-fiat pairs and 224 bitcoin-to-crypto pairs corresponding, respectively, to 348,529 and 340,681 observations. The bitcoin-to-crypto pairs in the final sample are bitcoin (BTC), bitcoin cash (BCH), ethereum classic (ETC), ethereum (ETH), litecoin (LTC), ripple (XRP), binance coin (BNB), cardano (ADA), doge coin (DOGE) and solana (SOL). We note that bitcoin discounts remain sizable despite restricting the raw data along several dimensions. Table 2 reports, for the cryptocurrencies in the sample, market capitalization, daily trading volume, and release dates. Bitcoin, ethereum, and ripple are among the main cryptocurrencies by market capitalization, while the remaining coins account only for a small fraction of the total crypto market capitalization. The bitcoin-to-fiat pairs in the final sample are the Australian dollar (AUD), the Brazilian real (BRL), the Canadian dollar (CAD), the Swiss franc (CHF), the Chilean peso (CLP), the Chinese yuan (CNY), the Czech Krone (CZK), the euro (EUR), the British pound (GBP), the Hong Kong dollar (HKD), the Indonesian rupiah (IDR), the Israeli new shekel (ILS), the Indian rupee (INR), the Japanese yen (JPY), the South Korean won (KRW), the Malaysian ringgit (MYR), the Nigerian naira (NGN), the Philippine peso (PHP), the Polish zloty (PNL), the Russian ruble (RUB), the Singapore dollar (SGD), the Thai bath (THB), the Ukrainian hryvnia (UAH), the U.S. dollar (USD), the Vietnamese dong (VND), the South African rand (ZAR). Note that we also adopt a broader definition for the geographical location of an exchange, which does not necessarily correspond to single sovereigns. Specifically, we define a geographical location as the set of exchanges located in a geographically or politically homogeneous location, like countries with the same, or similar, regulation (i.e., European countries). Specifically, after direct inspection of each exchange website to retrieve information about the legal location of the exchange, we could assign each exchange to one of the following locations: Africa, Australia, Canada, China, East Asia, Eastern Europe, European Union, the United Kingdom, India, Japan, South Korea, Latin America, Russia, Singapore, Turkey, and the U.S. East Asia includes Vietnam, Thailand, and Singapore; Latin

America includes Chile and Brazil; China also includes Hong Kong. Tables A2 and A3 in the Online Appendix provide additional information on the number of observations, and exchanges, for each currency and location.

Table 2: Cryptocurrencies

Rank	Name	Market Cap in bln $	Price in $	Volume in bln $ (24h)	Release Date	Supply Max in mln	Supply Circulating in mln
1	Bitcoin	583.7	30,035	14.1	9-Jan-09	21	19.4
2	Ethereum	228.7	1903	7.0	30-Jul-15	No cap	120
4	Ripple	39.1	0.74	2.4	26-Sep-13	39809.1	39.1
5	Binance Cash	37.7	242.1	0.68	14-July-2017	21	155.8
7	Cardano	10.7	0.30	0.30	27-Sep-2017	45	34.9
8	Solana	10.3	25.62	0.97	16-Mar-2020	511	402.6
9	Doge Coin	9.6	0.06	0.36	6-Dec-2013	No cap	140.2
12	Litecoin	6.7	91.26	0.58	7-Oct-11	84	73.3
18	Bitcoin Cash	4.6	236	0.37	1-Aug-17	21	19.4
28	Ethereum Classic	18.7	2.6	0.14	25-Oct-16	No cap	142.0
	of top 5000 crypto	1182		31.1			

Notes: This table lists the market capitalization, in billions of U.S. dollars, of the cryptocurrencies in our sample. For each currency, the table also reports its rank in terms of market capitalization, dollar price, daily volume in billions of U.S. dollars, date of release, and maximum and circulating supply. We do not include in our sample stable coins, like tether, which is the third cryptocurrency by market capitalization. Data are for July, 2023, from Coinmarketcap. The label "of top 5,000 crypto" refers to all the cryptocurrencies tracked by the data aggregator.

3 Bitcoin Discounts

In this section, we first establish stylized facts about bitcoin price dispersion at each point in time, and its variation over time, across geographical locations of the exchanges, and currency pairs. We then perform a variance decomposition into time, geographical location, and currency components, and provide estimates for the determinants of bitcoin discounts.

3.1 Measuring Bitcoin Discounts

We take the perspective of investors who can trade bitcoin at time t in a set of $m = 1, \dots, M$ markets (i.e., exchanges in different locations). We denote with $P^{\star}_{m,j,t}$ the units of currency $j = 1, \dots, J$ required to buy one bitcoin in market m at time t. We think of markets as cryptocurrency exchanges, like Kraken. We also denote with S^j_t the spot exchange rate expressed in units of currency j (i.e., fiat or crypto) per U.S. dollar, which we take as numeraire. Then, the U.S. dollar price of one bitcoin in market m and currency j at time t is:

$$P_{m,j,t} = \frac{P^{\star}_{m,j,t}}{S^{j}_{t}}.$$

In the absence of frictions, investors should get the same units of bitcoin per dollar in each market m and against each currency j. In practice, there exist large, persistent, and time-varying price differences. For example, at the end of 2017, the financial press, as well as investors' online forums, went to great length to discuss and analyze the so-called "Kimchi Premium", i.e., the fact that buying bitcoin on South Korean exchanges in Korean won was much more expensive than in other exchanges across the globe, after accounting for the currency conversion (for example, see Choi et al., 2022; Eom, 2021; Lee and Oh, 2022; Chen et al., 1993; Ok et al., 2023).[7] This is exemplified in the top left panel of Figure 1, which plots the U.S. dollar bitcoin price on three Korean exchanges (i.e., Bithumb, Korbit, and Coinone), along with the volume-weighted price across all other sample pairs. Starting around December 2017, bitcoin prices began to diverge, and by January 2018, the price on Korbit was 42 percent higher than the benchmark price. This is not an isolated instance: the top right panel of the same figure shows that, since 2020, the bitcoin price in Argentina has been almost twice as large as the benchmark price. Similarly, the bottom left panel of the figure indicates that, in more recent data, the bitcoin price in Nigeria has consistently remained above the benchmark price. Lastly, the bottom right panel demonstrates that the bitcoin price in Russia, after the start of the Ukraine war, has been higher than the benchmark price in some exchanges, such as Exmo.

In all four examples, the persistence of the substantial price differences can be attributed to capital controls. In Argentina, the bitcoin discounts reflect the stringent controls on the purchase of U.S. dollars, which are traditionally used as a store of value during periods of high inflation,

[7]The literature on the "Kimchi Premium" documents a positive association between the Kimchi premium, trading volume, and price volatility (Choi et al. (2022); Eom (2021)). It also examines the relationship between the Kimchi premium and market frictions (Lee and Oh (2022)). Chen et al. (2022) considers a behavioral explanation, while Ok et al. (2023) argues that the puzzle is still unresolved in the more recent data.

and the resulting in divergence between the official and market peso-to-dollar exchange rate. In Korea, bitcoin discounts are related to restrictions on capital flows, limiting the transfer of fiat currency in and out of the country. In Nigeria, bitcoin discounts are also linked to capital controls. Finally, in Russia, bitcoin discounts are connected to the international sanctions imposed after the invasion of Ukraine.[8]

Motivated by these examples, one might conclude that the exchange location is the only driver of the dispersion in bitcoin prices, as argued by the existing literature (see, for example, Makarov and Schoar (2020)). In this paper, we provide novel evidence that the exchange location is an important component, which accounts to approximately 50% of the overall variation in bitcoin prices for fiat pairs to less than 15% for crypto pairs. Further, we establish that additional components, like the exchange-quality and currency components, are quantitatively important.

In order to measure the dispersion in daily bitcoin prices, at each point in time, we introduce bitcoin discounts, defined as the ratio between the bitcoin dollar price in market m and currency j, and the volume-weighted average price across all markets and currencies, which we take as the benchmark price:

$$D_{m,j,t} = \frac{P_{m,j,t}}{\Psi_t} - 1 \tag{1}$$

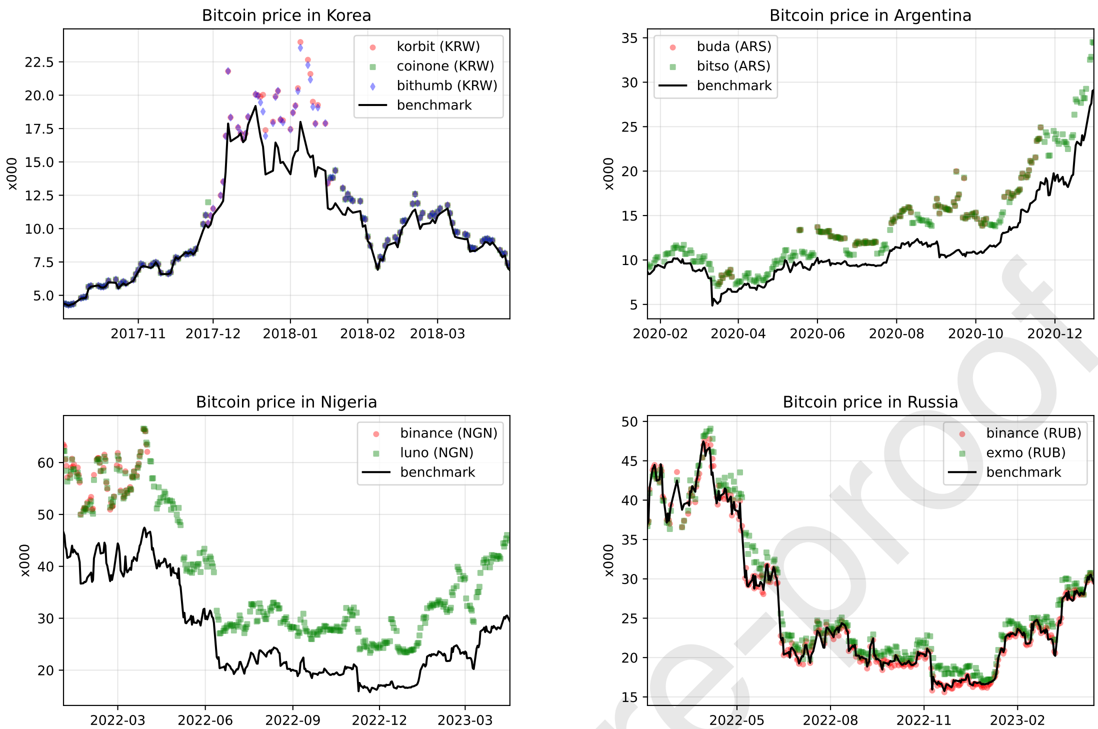
where

$$\Psi_t = \frac{\sum_m^M \sum_j^J P_{m,j,t} Q_{m,j,t}}{\sum_m^M \sum_j^J Q_{m,j,t}} \tag{2}$$

and $Q_{m,j,t}$ is the volume of transactions in market m, currency j, and date t, expressed in bitcoin. If $D_{m,j} < 0$, then bitcoin is relative "cheap" in market m and currency j and, for each dollar, investors get more bitcoin than when using the benchmark price. On the contrary, if

[8] For more details, see, e.g., https://www.bloomberg.com/news/articles/2021-11-24/a-huge-arbitrage-opportunity-has-just-opened-up-in-crypto for Korea; https://www.bloomberg.com/news/articles/2019-08-13/bitcoin-draws-premium-in-argentina-and-hong-kong-amid-sell-off for Argentina; https://cointelegraph.com/news/bitcoin-premium-hits-60-in-nigeria-as-it-limits-atm-cash-withdrawals for Nigeria; https://finance.yahoo.com/news/bitcoin-premium-emerges-ukrainian-markets-174004872.html for Russia.

Figure 1: Four Examples of Bitcoin Discounts



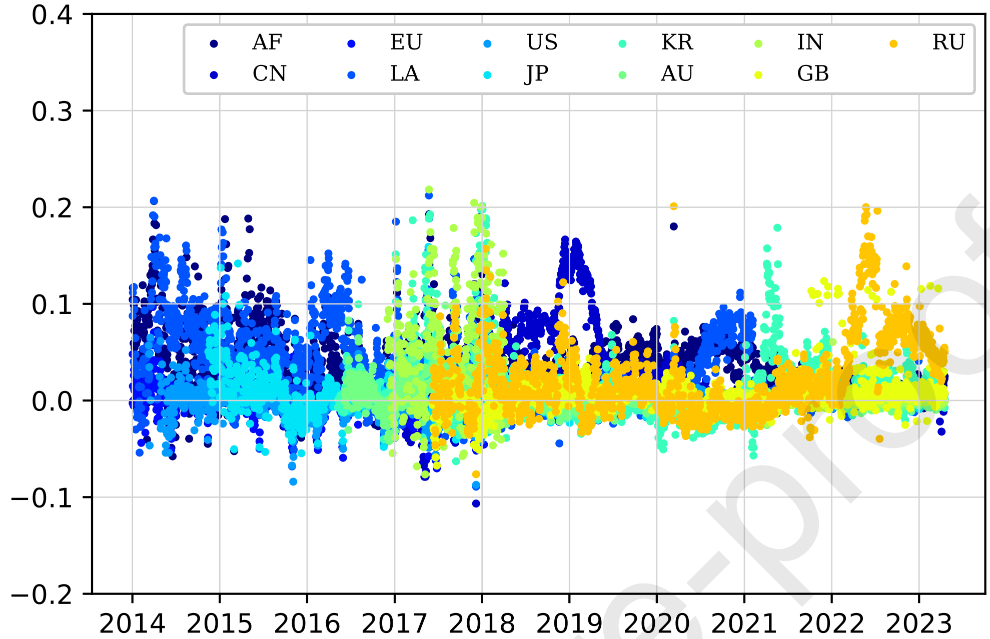
Notes: The left top panel plots the daily U.S. dollar prices for three Korean won (KRW) pairs on the Bithumb, Korbit, and Coinone, respectively. The right top panel plots the daily U.S. dollar prices for two Argentinian peso (ARS) pairs on Bitso and Buda. The bottom left panel plots the daily U.S. dollar prices for two Nigerian Naira (NGN) pairs on Binance and Luno. The bottom right panel plots the daily U.S. dollar prices for two Russia Ruble (RUB) pairs on Binance and Exmo. The black line denotes the benchmark volume-weighted price across all other sample pairs. Data are daily from Cryptocompare (bitcoin prices) and Thomson Reuters (spot exchange rates).

$D_{m,j} > 0$, then bitcoin is relative "expensive" in market m and currency j, and investors get less bitcoin than when using the benchmark price. When $D_{m,j} = 0$, then investors get the same number of bitcoin per U.S. dollar in all markets and currencies. While the "Kimchi Premium" or the bitcoin to Argentine peso are just illustrative examples, Figure 2 presents a snapshot of the evolution of bitcoin discounts across all locations. Specifically, we plot the time-series of daily bitcoin discounts using different colors for the different geographical locations associated with the exchanges. The figure reveals a simple stylized fact: discounts are typically different from zero, positive or negative, and very volatile for all locations. In fact, daily discounts can be as large as 40% in absolute value, and a casual inspection of the figure also reveals that discounts are persistent. A simple first-order autoregressive model explains a large fraction of the time-series variability of bitcoin discounts for most currency pairs. For example, Figure 3 reveals that for all pairs, the distribution of autoregressive coefficients is skewed to the right with a large mass in the right tail, for values of the autoregressive coefficients greater than the mean value of 0.43. Although these stylized facts are qualitatively similar for both bitcoin-to-fiat and bitcoin-to-crypto pairs, discounts for the fiat pairs are more persistent. In fact, the mean autoregressive coefficients are equal to 0.59 for fiat pairs, and to 0.41 for crypto pairs (see Table 3). In what follows, we establish that a fraction of the variability in discounts, across pairs and exchange locations, and over time, depend on market frictions that limit arbitrage-like activity by investors. Frictions are likely to be larger for fiat pairs. For example, in the case of bitcoin-to-fiat pairs, investors must use the currency spot market, which is open only during business days and hours, and the settlement of transactions can take up to several days. Besides, the transfer of capital across exchanges located in different markets may be subject to country-specific regulations, or controls, that could further delay the process. In contrast, for the case of bitcoin-to-crypto pairs, the settlement of transactions requires a significantly shorter amount of time (usually measured in hours and not days), and the transfer of currency is not effectively subject to country-specific regulations and capital controls.[9]

We compute statistics for the distribution of bitcoin (gross) discounts over every market m, currency j, and day t, and report them in Table 3. When we consider all pairs, the average standard deviation is 3.9%; the mean 90-10 percentile ratio is 1.064; the mean 90-50 percentile ratio is 1.035; and the mean 50-10 percentile ratio is 1.028. These numbers reveal substantial dispersion in the bitcoin prices in a given market, currency, and day. We find a similar price dispersion also when we consider only bitcoin-to-fiat or bitcoin-to-crypto currency pairs. For fiat pairs, the standard deviation is higher and equal to 5.3%. In contrast, for crypto pairs, it is lower and equal to 2.4%. The values for the kurtosis, which are large for both fiat and crypto pairs, indicate tail data exceeding the tails of the normal distribution. Finally, discounts are large: for fiat pairs, they range from -43% to 99% and for crypto pairs from -23% to 30%. As a benchmark, consider that Gagnon and Karolyi (2010) find that the mean discount for ADRs is 4.9 basis points with a daily standard deviation

[9] In the Online Appendix, we show that discounts for fiat pairs are larger, more volatile, and more persistent than for crypto pairs (see Figure A2 and Figure A4).

Figure 2: Discounts Across Locations



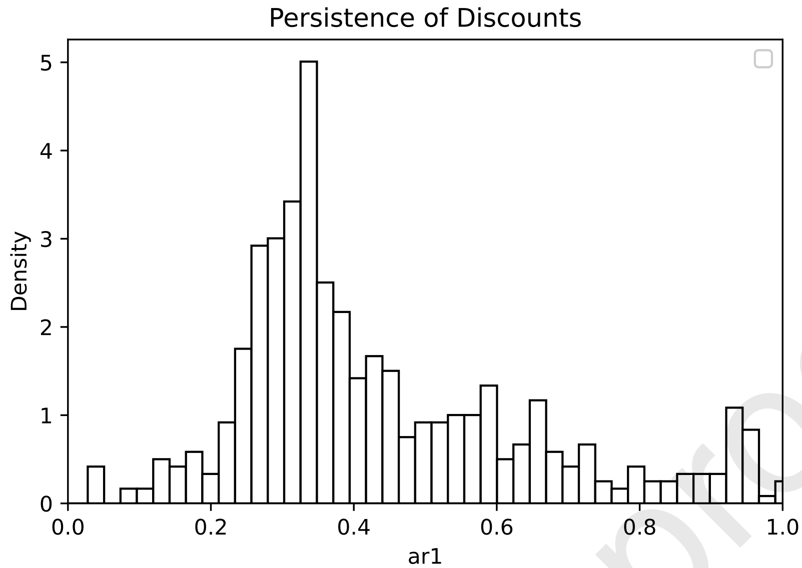
Notes: This figure plots the bitcoin discounts for all the bitcoin-to-fiat and bitcoin-to-crypto pairs in our sample. Different colors correspond to different geographical locations associated with the exchanges. Discounts are defined according to Equation (1). Data are daily from Cryptocompare and Thomson Reuters for the period 1/1/2014–4/18/2023.

for a given stock pair of 1.4%, even though they observe extreme deviations as large as -40% and 127%. Wang and Jiang (2004) find average daily discounts for H-shares in Hong Kong relative to A-shares in mainland China of the same Chinese companies of 75%. Du et al. (2018) find mean daily deviations for the covered interest parity that range from 6 to 19 basis points annualized, with standard deviations from 4 to 23 basis points.

3.2 Deconstructing Bitcoin Discounts

Before presenting the results of the decomposition of bitcoin discounts, it is convenient to start with an illustrative example. We start by considering one randomly picked day, that is, Wednesday June 20 2018, and a sample containing only bitcoin-to-dollar pairs across different exchanges in dif-

Figure 3: Persistence of Discounts (all pairs)



Notes: This figure presents the distribution of the autoregressive coefficients obtained by estimating an autoregressive model of the first order for bitcoin discounts and all pairs. Discounts are defined according to Equation (1). Data are daily from Cryptocompare and Thomson Reuters for the period 1/1/2014–4/18/2023.

Table 3: Average Statistics of Price Distribution

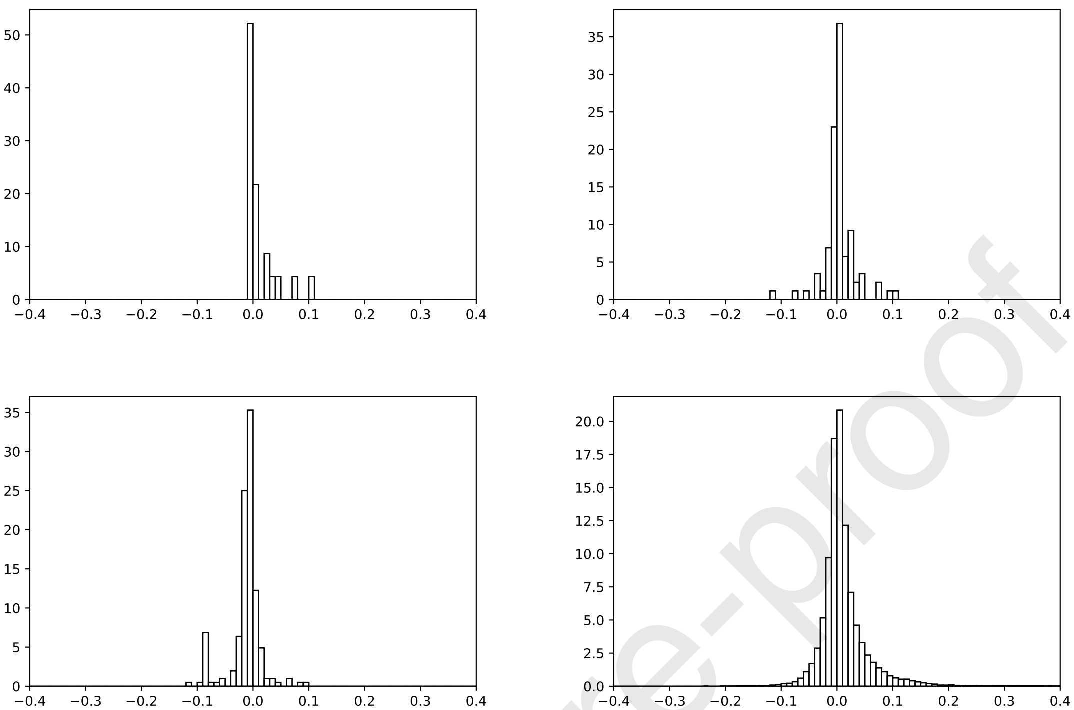
Data	Standard Deviation	Minimum	Maximum	90-10 ratio	50-10 ratio	90-50 ratio	Skewness	Kurtosis	AR(1)
fiat	0.053	0.571	1.994	1.052	1.017	1.035	7.128	80.249	0.586
crypto	0.024	0.771	1.295	1.047	1.022	1.024	0.429	9.032	0.416
total	0.039	0.757	1.848	1.064	1.028	1.035	4.401	57.268	0.436

Notes: The table reports average statistics for bitcoin (gross) discounts, defined as $(1 + D_{m,j,t})$. The first row refers to bitcoin-to-fiat pairs; the second row to bitcoin-to-crypto pairs; the last row to all pairs. All statistics are volume weighted across currencies, exchanges, and days. Note that we apply a small winsorization to the data (i.e., 0.1% at the top and bottom of the distribution) which explains, for example, why the minimum discount is larger for the crypto pairs than for all pairs. Data are from Cryptocompare and Datastream for the period 1/1/2014–4/18/2023.

ferent locations. For this specific date, our sample contains only 23 observations, all corresponding to bitcoin-to-dollar pairs traded on 33 different exchanges, and, from the perspective of dollar-based investors, there is no exchange rate risk. We compute bitcoin discounts using Equation (1) and plot in the top left quadrant of Figure 4 their distribution. The standard deviation of discounts is 2.7%, and discounts range from -0.7% to 10.1%. Next, we enlarge the sample by including all the bitcoin-to-fiat pairs available on the same day. We now have 87 observations. The standard deviation increases to 2.8%, and discounts now range from -11.% to 10.1%. When we additionally include in the sample also all the bitcoin-to-crypto pairs, we arrive at 174 observations. The standard deviation drops to 2.7%, and discounts range from -11.4% to 9.3%. Finally, we consider the full length of the sample, from January 1 2014, to April 18 2023. In this case, the standard deviation increases to 4.7%, and discounts range from -24.3% to 66.1%. The remaining quadrants of Figure 4 plot the distribution of discounts for these last three steps. This example reveals that discounts are large, on a given day, for both fiat and crypto currency pairs. Furthermore, the example shows that the heterogeneity in price dispersion is primarily due to variation over time in relative bitcoin prices. This is exemplified by the evolution of the cross-sectional variance of discounts reported in the top left panel of Figure 5, which reveals that price dispersion varies greatly over time for both fiat and crypto pairs, with the cross-sectional standard deviation ranging from 0.1% to 12% for fiat and 0% to 8% for crypto pairs, and that it did not decline in the more recent sample. Note that as the number of fiat and crypto pairs in our sample increases and then declines (top right panel of Figure 5), the average exchange "quality" gap between fiat and crypto pairs narrows. This evidence motivates our inclusion, in the next section, of measures of counter-party risk in the analysis of the determinants of discounts.

For the documented large and highly dispersed discounts to exist and persist, there must exist frictions and restrictions to trade that may have originated from two conceptually different sources. First, there could be heterogeneity in bitcoin prices across the different exchange locations, for example, because of differences in regulation, transparency, and capital controls. In this case, bitcoin trades at different prices because some pairs are traded in a relatively "cheap" location, while others are traded in a relatively "expensive" location. Second, there could be heterogeneity in bitcoin prices for different pairs traded on exchanges in the same location. In this case, bitcoin trades at different prices because it corresponds to different currency-pairs. While frictions (e.g., delays in the speed of execution and capital controls.) are necessary to explain the existence of the discounts, some heterogeneity in the demand and supply for bitcoin in the different markets is required to explain the emergence of discounts in the first place (see, e.g., Cochrane (2002)).

Figure 4: Deconstructing Bitcoin Discounts



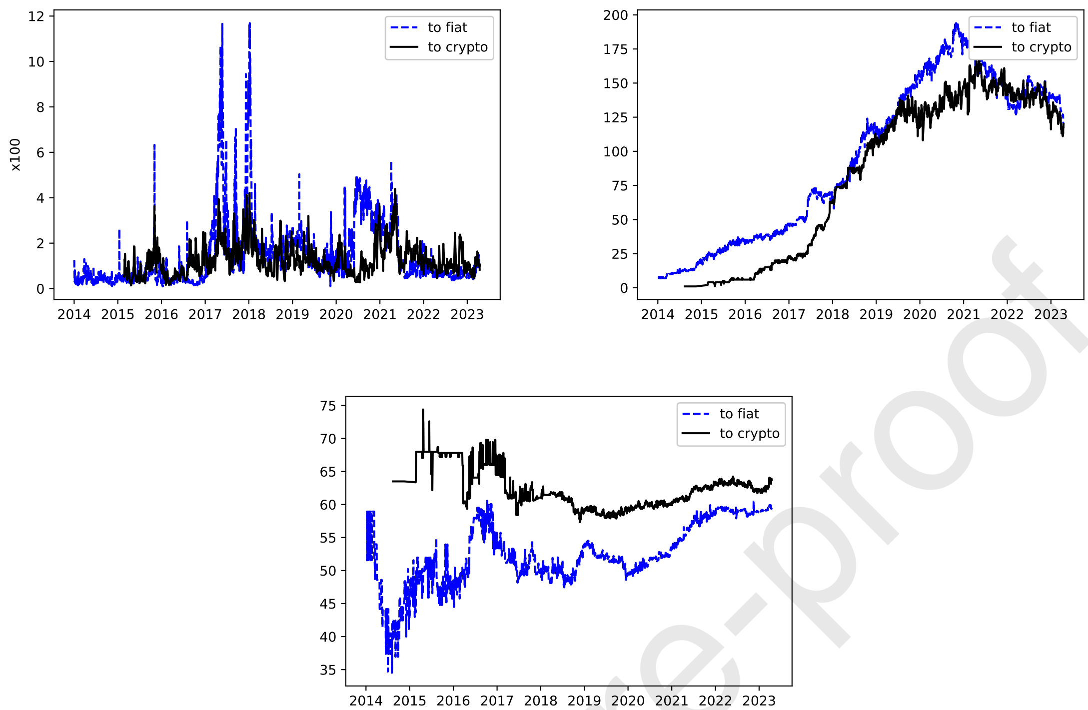
Notes: The figure shows the distribution of bitcoin discounts for different samples. The subplots "USD", "Fiat" and "Crypto" are for a single day, i.e., Wednesday, June 201 2018 and correspond, respectively, to all the bitcoin-to-dollar pairs, all the bitcoin-to-fiat pairs, and all the bitcoin-to-fiat and bitcoin-to-crypto pairs. The subplot "Overall" corresponds to all bitcoin pairs over the full length of the sample. Data are daily from Cryptocompare and Thomson Reuters for the period 1/1/2014–4/18/2023.

We formalize a decomposition of price dispersion, building on Kaplan and Menzio (2015), and start with introducing additional notation. First, for exchange location $g = 1, \dots, G$ and time t, $\Psi_{g,t}$ is the volume-weighted average price of all pairs with location g:

$$\Psi_{g,t} \equiv \mathbf{E}_{m,j}\left[P_{m,j,t}|g\right] = \frac{\sum\limits_{m\in\Omega_g}\sum\limits_{j=j(m)} P_{m,j,t}Q_{m,j,t}}{\sum\limits_{m\in\Omega_g}\sum\limits_{j=j(m)} Q_{m,j,t}} \tag{3}$$

where $m \in \Omega_g$ indicates all markets in the set Ω_g, which contains all markets with location g; $j = j(m)$ denotes all pairs traded in market m; $\mathbf{E}_x(Y_{x,z}|z)$ denotes the expectation conditional on z across the x dimension. Second, for location g, currency j and time t, $\Psi_{g,j,t}$ denotes the

Figure 5: Cross-sectional Variability of Bitcoin Discounts



Notes: The figure plots the evolution of the cross-sectional standard deviation in bitcoin log discounts (top left panel); of the number of available pairs (top right panel), and of the average exchange quality (bottom center panel), for fiat (blue dashed line) and crypto (black solid line) pairs. Data are daily from Cryptocompare.com and Datastream for the period 1/1/2014–4/18/2023. Exchange quality is measured in grade points. "High" quality exchanges have a grade point of 45 or higher.

volume-weighted average price of all the j pairs traded in markets with location g:

$$\Psi_{g,j,t} \equiv \mathbf{E}_m \left[P_{m,j,t} | g, j \right] = \frac{\sum_{m \in \Omega_g} P_{m,j,t} Q_{m,j,t}}{\sum_{m \in \Omega_g} Q_{m,j,t}} \tag{4}$$

Finally, $\Psi_{g,q,t}$ and $\Psi_{g,q,j,t}$ denote, respectively, the volume-weighted average prices in (3)-(4) for high- or low-quality (q) exchanges.

The data allow us to decompose log discounts ($d_{m,j,t} = \log(1 + D_{m,j,t})$), in market m, currency j and day t, into four components:

$$d_{m,j,t} \equiv p_{m,j,t} - \psi_t = (\psi_{g,t} - \psi_t) + (\psi_{g,q,t} - \psi_{g,t}) + (\psi_{g,q,j,t} - \psi_{g,q,t}) + (p_{m,j,t} - \psi_{g,q,j,t}), \tag{5}$$

where lower case variables denote logged variables (i.e., $p = \log P, \psi = \log \Psi$), and ψ_t is the log of the volume-weighted average price across all pairs and markets. The first term in Equation (5) is the mean relative price of all pairs traded in location g with respect to the mean bitcoin price. The second component is the mean relative price of all pairs traded in exchanges of quality q in location g with respect to the mean price of all pairs traded in the same location. The third component is the mean relative price of all currencies j traded in exchanges with quality q in location g with respect to the mean price of all pairs traded in exchanges of the same quality and location. Finally, the fourth component is the relative price of each pair with respect to the mean price of the same currency j in exchanges of quality q in location g. By applying the law of total variance conditioning on g, q and j we obtain:

$$\mathbf{Var}_{m,j}[d_{m,j,t}] = \mathbf{Var}_g\left[\mathbf{E}_{m,j}\left[d_{m,j,t}|g\right]\right] + \mathbf{E}_g\left[\mathbf{Var}_{m,j}\left[d_{m,j,t}|g\right]\right] \tag{6}$$

$$= \mathbf{Var}_g[\psi_{g,t} - \psi_t] + \mathbf{E}_g\mathbf{Var}_q[\psi_{g,q,t} - \psi_{g,t}] +$$

$$+ \mathbf{E}_{g,q}\mathbf{Var}_j[\psi_{g,q,j,t} - \psi_{g,q,t}] + \mathbf{E}_{g,q,j}\mathbf{Var}_m[p_{m,j,t} - \psi_{g,q,j,t}], \tag{7}$$

where $\mathbf{Var}_x(Y_{x,z}|z)$ is the variance conditional on z across the x dimension. Equation (6) refers to the decomposition only with respect to location. In this case, the first component captures the dispersion in the conditional means of log discounts in different locations; while the second component captures the mean dispersion of log discounts in different locations. In Equation(7) we further decompose the second term of Equation (6) into three components: the first captures the mean price dispersion between exchanges of high- and low-quality conditional on location; the second captures the mean dispersion across different currencies conditional on location and quality; the last captures the residual variance.

Table 4 in this response, presents the results of the variance decomposition for all fiat (panel A) and crypto pairs (panel B), based on Equation 7 in the paper. On average, we find that the dispersion in prices across locations is the main component of the variability of discounts for fiat pairs, less so for crypto pairs. Specifically, the location component explains approximately

50% of the total variation for fiat pairs and 11.2% for crypto pairs. Dispersion in prices between high- and low-quality exchanges, within a location, explains approximately 9-7% for all pairs, but it gained recently importance for fiat pairs accounting for up to 92.9%. For fiat pairs, 9.8% of the total variation is accounted for by the dispersion in prices across different currencies, and the residual component accounts for 31.1%. For crypto pairs, 74.4% of the total variation is accounted for by the dispersion in prices across different currencies, while the residual component is smaller and equal to 8.5%. The large differences between the minimum and maximum shares explained by each components indicate time-variation in the decomposition (see the last two columns of Table 4). Figure 6 plots the evolution of these shares over time and shows that, for both fiat and crypto pairs, the contribution of the location component (red area) has been declining over time. For fiat pairs, the contribution of "quality" and "currency" gained relevance, while for crypto pairs the relative contribution of the currency component is larger.

Table 4: Variance Decomposition

Panel A: bitcoin-to-fiat

	mean	std	min	max
location	0.503	0.230	0.045	0.962
quality	0.089	0.092	0.003	0.413
currency	0.098	0.079	0.009	0.352
residual	0.310	0.261	0.000	0.908

Panel B: bitcoin-to-crypto

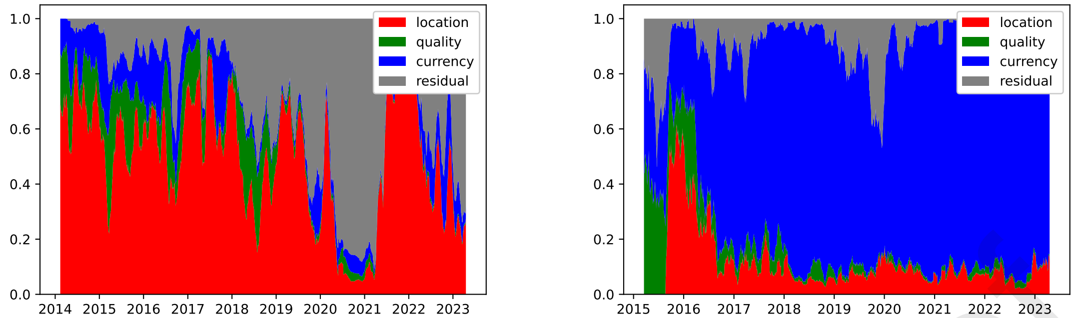
	mean	std	min	max
location	0.110	0.110	0.000	0.619
quality	0.055	0.084	0.003	0.522
currency	0.772	0.172	0.127	0.955
residual	0.081	0.095	0.002	0.469

Notes: The table reports results of the variance decomposition of bitcoin discounts. The columns report, for each components in Equation (7), the mean share of explained variance; the standard deviation; the minimum and maximum values. Panel A refers to bitcoin-to-fiat pairs, Panel B to bitcoin-to-crypto pairs. Numbers are reported in percentage and smoothed using a 30 days window. Data are daily from Cryptocompare, Datastream, SimilarWeb and Alexa for the period 1/1/2014–4/18/2023.

3.3 Derivative trading and cryptomarket efficiency

The documented existence of large and persistent bitcoin discounts questions the efficiency of cryptomarkets. In this section we conduct a formal examination of the potential impact of the launch of Bitcoin futures on the price efficiency of the Bitcoin spot market. This analysis stems from a well-established literature that establishes a connection between market completeness

Figure 6: Variance Decomposition of Bitcoin Discounts



Notes: The figure plots the time-series of the variance decomposition of bitcoin discounts following Equation (7). We consider four components: location (red); currency (blue); quality (green), and residual (black). The left panel refers to fiat pairs, the right panel to crypto pairs. Shares are smoothed using a 30 days window. Data are daily from Cryptocompare.com, Datastream, SimilarWeb and Alexa for the period 1/1/2014–4/18/2023.

and market efficiency. Notably, Figlewski and Webb (1993) demonstrate how options trading contributes to both transactional and informational efficiency in the stock market, while Pan and Poteshman (2006) reveal that the price of equity derivatives carries information regarding future equity prices. In the literature on cryptomarkets, Corbet et al. (2018) present an early study that investigates the effects of bitcoin futures' introduction. Their findings indicate an increase in spot volatility following the launch of bitcoin futures. Additionally, Kapar and Olmo (2019) argue that price discovery predominantly occurs in the futures market, while Augustin et al. (2020), using a more recent sample, discover that the introduction of bitcoin futures contracts has led to heightened price synchronicity for bitcoin-to-dollar pairs across exchanges. They interpret this as evidence of reduced frictions that impede arbitrage activity.

To assess whether the introduction of futures contracts has contributed to the price efficiency of the Bitcoin spot market, we empirically test the hypothesis that the average variance in bitcoin discounts, across pairs, decreased subsequent to their launch. We conduct these tests using various time windows surrounding the introduction dates, specifically the 10th of December 2017 and the 18th of December 2017. The former corresponds to the introduction of CBOE Futures, while the latter aligns with the introduction of CME Futures (e.g., refer to Corbet et al. (2018)). Augustin et al. (2020) documents how the contracts launch was largely unanticipated, supporting a causal interpretation of the effect of futures introduction which we test.[10]

Results are presented in Table 5, which reveals that we reject the null of a decrease in the variance of bitcoin discounts for bitcoin-to-fiat pairs for both dates and all windows. For bitcoin-to-crypto pairs, we reject the null in the case of the date of launch of CBOE Futures, regardless of the window. In contrast, in the case of the date of launch of CME futures, we cannot reject the null of a decrease in the variance for both the 3-month and 1-month window.

Table A10 in the Online Appendix presents results of a more granular test at the location level which reveals a heterogeneous effect on price efficiency at the location level of the launch of derivatives trading.

Table 5: The impact of the launch of derivative trading on cryptomarket efficiency

sample	date	full	3M window	1M window
fiat	10/12/2017	-11.14	-6.49	-6.00
fiat	18/12/2017	-10.81	-6.35	-10.67
crypto	10/12/2017	5.09	-3.28	-4.43
crypto	18/12/2017	6.22	0.58	-1.33

Notes: The table presents the results of a t-test of the hypothesis that the average variance of bitcoin discounts decreased after the launch of bitcoin futures contracts. The launch dates are the 10th of December 2017 (introduction of CBOE Futures) and the 18th of December 2017 (introduction of CME Futures). We consider a window based on the full sample before and after the launch dates, as well as a 3 and 1 month window. The reported values are t-stats. Data are daily from Cryptocompare.com, and Thomson Reuters for the period 1/1/2014–4/18/2023.

4 Explaining Bitcoin Discounts

In this section, we estimate time-series and panel regressions to investigate the contribution of observable factors to the variation in bitcoin discounts over time, and across different markets and currencies. We consider a large set of candidate factors and group them into four categories: counter-party risks; liquidity risk; blockchain factors; and crypto factors. While in the time-series regressions, we consider aggregate factors (e.g., the mean bid/ask spread across all markets and

[10]We refer the interested reader to Augustin et al. (2020) for institutional details about the CBOE and CME bitcoin futures contracts.

pairs), in the panel regressions, when possible, we consider location-specific factors (e.g., the mean bid/ask spread across all pairs in a given location). The empirical analysis is based on a shorter sample, ending at the beginning of 2020, due to the data availability.

4.1 Candidate Factors

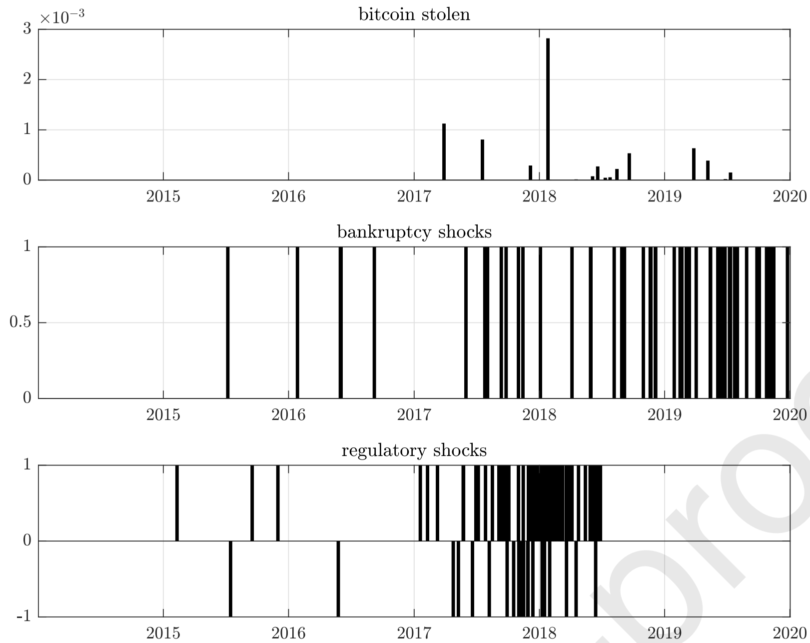
We start by describing the construction of the factors and by motivating their relation with bitcoin discounts. Section A, in the Online Appendix, provides additional details on the factors, on their construction, and the data sources.

Counter-party risks. Cryptocurrency exchanges function in many ways like brokers or banks. Customers buy and sell bitcoins (or other cryptocurrencies), but typically maintain a balance of both fiat currency and bitcoin on the exchanges without retaining direct access to the currency (see, for example, Gandal et al. (2018)). All the trades on a given exchange are completed *off* the blockchain. When investors deposit bitcoin on an exchange, these are put in a shared account, called a "wallet", that the exchange controls (i.e., these bitcoins are like deposits for a bank). While the exchange keeps track of investors' balances and of all the transactions, bitcoin transactions are recorded on the blockchain only at the time of transfer to and from the exchange. When investors withdraw coins, then the blockchain is informed, and bitcoins are transferred to the investors' personal wallets. Therefore, trading bitcoin on exchanges is similar to holding IOUs and involves the risk that one party (i.e., the exchange) defaults on the transaction. For example, when in February 2014, Mt. Gox, the largest exchange by trading volume at the time, declared bankruptcy, approximately 850,000 bitcoins belonging to customers were stolen (450 million U.S. dollars at the time). While Mt. Gox failure is probably the most extreme example of counter-party risk after a security breach, there are many more examples of temporary and permanent exchange shut-downs. Moore and Christin (2013) find that, by early 2013, 45% of Bitcoin exchanges had closed, and many of the remaining markets were subject to frequent outages and security breaches. Vasek and Moore (2015) investigate denial-of-service attacks against cryptocurrency exchanges and document 58 such attacks. Makarov and Schoar (2022) consider exchange hacks in more recent data.[11] In order to capture counter-party risks, we consider the following measures. The first measure is based on the amount of bitcoin stolen as a fraction of the total bitcoin supply. In our sample, the largest amount stolen is approximately equal to 0.28 percent of the total supply of bitcoin and occurred on January 26, 2018, when Japan's largest cryptocurrency exchange, Coincheck, was hacked. The theft caused Coincheck to suspend trading indefinitely. The second measure is based on a bankruptcy indicator. Specifically, we build a dummy variable that takes a value of one in the days when at least one of the exchanges in our sample permanently shuts down. For equity markets, Nakamura et al. (2013) show that they tend to be closed precisely at times when expected returns should be low. The third measure is based on a regulatory shock indicator, based on data collected by Auer and Claessens (2018), and manually integrated with data from Cointelegram, which takes the value of *one* for new crypto-friendly regulation, and of *negative one* for new regulation that restricts the trading, use, or transfer to cryptocurrency. Note that the latter can also be interpreted as an indicator of political or regulatory risk. For example, new domestic regulation could shut down exchanges, or set capital controls, as in China in 2017 (see Borri and Shakhnov, 2019). Figure 7 summarizes the properties of these three measures. In addition, Figure A1 plots the evolution of the daily fraction of inactive pairs, i.e. pairs with the daily volume equals to zero.

Liquidity risks. In order to capture liquidity risks, we build two different measures. The first measure is based on trading volume. Specifically, we consider the daily mean trading volume in bitcoin, normalized by total bitcoin supply, across all exchange-currency pairs. The mean trading volume is approximately 0.08 percent of the total bitcoin supply or about 12,000 bitcoin per day. After a peak at the beginning of January 2017, when it was approximately equal to 1.8 percent of the bitcoin supply, the trading volume dropped substantially and has averaged around 0.01 percent since then. Borri and Shakhnov (2019) show that the large drop in volume is mostly explained by the "China shock", a dramatic and unexpected change in regulations by the Chinese authorities

[11] In Table A6 in the Online Appendix we provide a list of critical issues involving some of the leading cryptocurrency exchanges in our sample.

Figure 7: Counterparty Risks



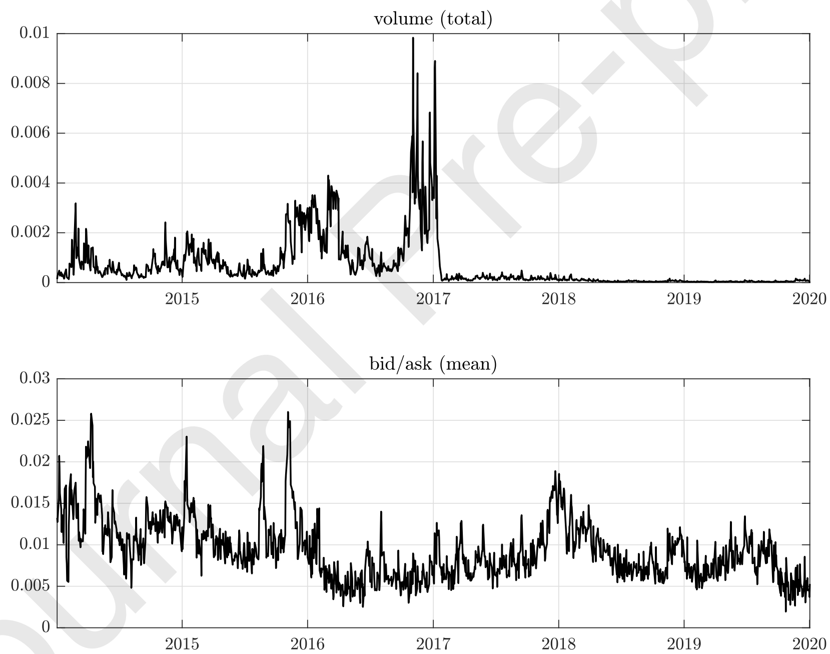
Notes: This figure plots three measures of cryptocurrency counter-party risks: the amount of bitcoin stolen in exchange hacks, expressed in percentage of total bitcoin market capitalization; a bankruptcy shock indicator that takes a value of one in the days when at least one of the exchanges in our sample permanently shuts down; a regulatory shock indicator that takes the value of one for new crypto-friendly regulation, and of negative one for new regulation that restricts the trading, use, or transfer to cryptocurrencies. Data are daily from Cryptocompare.com, walletexplorer.com, Auer and Claessens (2018), Cointelegram, https://www.hackmageddon.com/, bitinfocharts.com, and Thomson Reuters for the period 1/1/2014–1/1/2020.

that severely restricted bitcoin trading in China. Note that our sample does not contain exchanges where only crypto pairs are traded. The fact that these are mostly unregulated has recently raised concerns about their reliability and reported trading volume (Bitwise, 2019). The second measure is based on bid/ask spreads. Specifically, we consider the daily mean bid/ask spreads across all exchange-currency pairs. We obtain daily bid/ask spreads data for a subset of 33 fiat-to-bitcoin pairs in our sample, and then estimate bid and ask prices for all fiat pairs using the predicted values from the following panel regression.

$$BA_{m,j,t} = \alpha_j + \gamma_t + B(L)v_{m,j,t} + A(L)hl_{m,j,t} + \epsilon_{m,j,t}, \tag{8}$$

where $B(L)$ and $A(L)$ are fifth-order lag polynomial, v is the log trading volume in bitcoins, hl is the lag of the high-low spread, and α_j and γ_t are currency and time fixed effects.[12] Instead, we do not have data on bid/ask spreads for bitcoin-to-crypto pairs. In our sample, the average bid/ask spread is approximately 1 percent, and we observe a large heterogeneity across exchange and currency pairs. For example, the mean cross-sectional volatility is almost as large as the sample mean and, specifically, around 0.8 percent. Although the bid/ask spread is currently only a few basis points for the most traded exchange-currency pairs, the progressive entry of new pairs and opening of new exchanges explain both the relatively high average bid/ask spreads and the large cross-sectional volatility. Figure 8 summarizes the properties of these two measures.

Figure 8: Liquidity

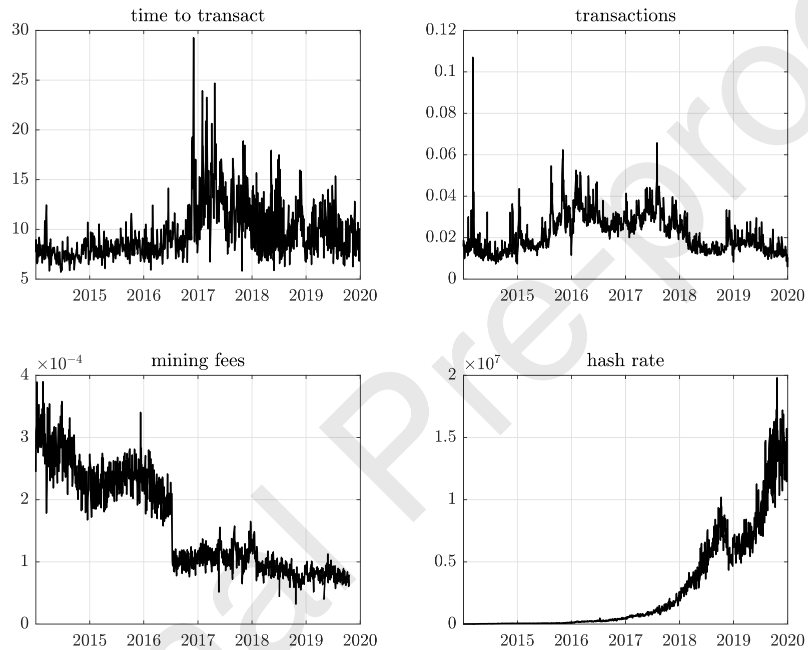


Notes: This figure plots the daily mean trading volume and bid/ask spreads across exchange and currency pairs. We divide the volume of transaction by the total bitcoin market capitalization. Data are daily from blockchain, Cryptocompare.com, Bitcoinity, and Thomson Reuters for the period 1/1/2014–1/1/2020.

[12] The (within) R-squared of our panel estimation is 20%, and most of the coefficients are significant at standard confidence levels. Specifically, higher volume and lower high-low spread are associated with smaller bid/ask spreads. In the Online Appendix, we estimate bid/ask spreads with alternative estimators based on daily high, low, and close prices. Note that while the bid-ask spreads for the most liquid pairs in our sample are small and consistent with other studies (Dyhrberg et al., 2018), those for the less liquid pairs are an order of magnitude larger. Finally, note that some exchanges do not offer a limit order book, but only matching of buy/sell orders. For these exchanges, our bid/ask estimates are a proxy of liquidity, as they do not explicitly post bid/ask prices.

Blockchain factors. The value of bitcoin, and cryptocurrency more in general, depends on fundamental characteristics of the blockchain. For example, Pagnotta and Buraschi (2018) and Biais et al. (2019) argue that the computing power and network size are related to the security and network benefits of the blockchain. We consider four blockchain factors. The first measure is the total value of transactions on the blockchain, expressed in bitcoin, and as a fraction of total bitcoin supply. The daily value of transactions has averaged around 2.2% of bitcoin supply, or 35,000 bitcoin per day, with a daily volatility of approximately 0.94% (or 15,000 bitcoin per day). The second factor is the median time-to-transact. The proof-of-work, required by the bitcoin blockchain to transfer bitcoins across exchanges, depends on the solution of a computationally challenging problem, which takes more time depending on the traffic on the network. The median time-to-transact captures the time to be accepted into a mined block and added to the public ledger. While the median transaction time has been roughly stable around 8 minutes throughout the sample, the average transaction time is much higher, and reached high values of about 11,000 minutes in January 2018 (when the median time was only 12 minutes). Note that exchanges typically require multiple confirmations before crediting a customer's account. For example, Kraken requires 6 confirmations for bitcoin, which correspond to approximately 1 hour. The third factor is a measure of mining activity, which combines block reward and transaction fees. A block reward refers to the number of bitcoins miners receive if they successfully mine a block of currency. The amount of the reward halves every 210,000 blocks, or roughly every four years, and is expected to hit zero around 2140. The current reward is equal to 6.25 bitcoin, and miners currently mine an average of 144 blocks per day. The halving in July 2016 explains the large drop in mining fees in Figure 9, while the next halving will occur in May 2024 (the halving of May 2020 is not in our sample). Transaction fees are the compensation for validating transactions, for adding them to the blockchain, and the exact amount of fees depends on network conditions and the data size of transactions. While transaction fees explain most of the volatility in our measure, block rewards explain its level. Lehar and Parlour (2020) further document high variation of Bitcoin fees, not only over time, but also within blocks. The fourth is the bitcoin hash rate, which is the measuring unit of the bitcoin network's processing power and is also a key security metric. The greater is the hashing power in the network, the greater is its security. The hash rate is expressed in terahashes, where one hash refers to one function solved by a computer (1 terahash = 10^{12} hashes). The hashing power is estimated from the number of blocks being mined and the current block difficulty, and its volatility also depends on the randomness of block discovery.[13] Figure 10 summarizes the properties of these four measures.

Figure 9: Blockchain Factors



Notes: This figure plots the daily median time-to-transact in minutes; a measure of mining fees, equal to the sum of transaction fees and block rewards; the hash rate; and the total transaction value on the blockchain in bitcoin. The hash rate is expressed in terahashes (1 terahash = 10^12 hashes). Data are daily from blockchain and coinmetrics.io for the period 1/1/2014–1/1/2020.

Crypto factors. We consider four factors related to the demand and supply on cryptocurrency markets. The first measure is simply the bitcoin price, which is a measure of the relative demand and supply of cryptocurrency. The bitcoin price has experienced a dramatic increase from around 800 U.S. dollars at the beginning of the sample, to a peak value of more than 18,000 U.S. dollars

[13] The hashing power is estimated as follows. Given the average time between mined blocks (T) and a difficulty (D), the hash rate per second is $H = 2^{32}D/T$ (see, for example, blockchain.com).

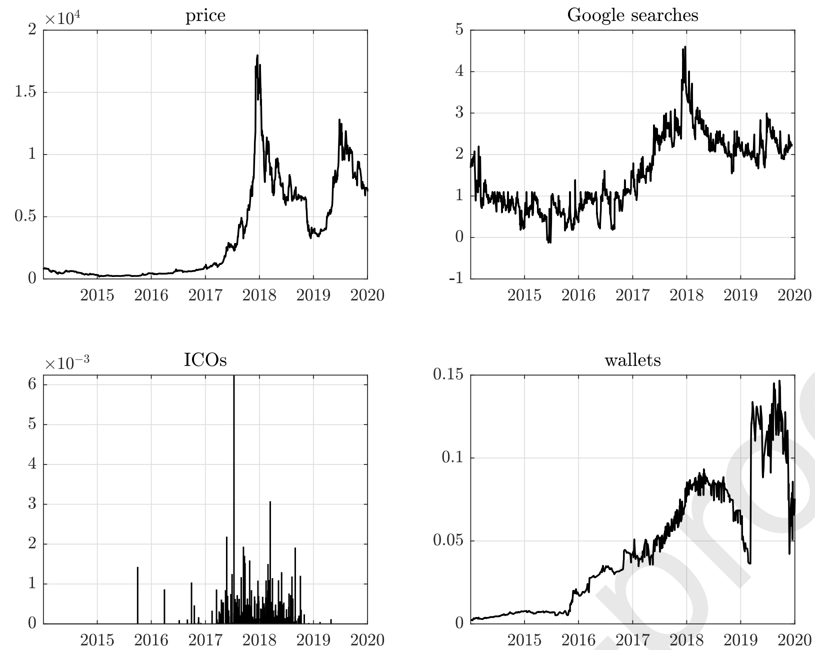
in December 2017. Since its peak, the bitcoin price has hovered around 7,500 U.S. dollars. The second measure is a proxy of demand, or sentiment, using Google Trends data, which capture the popularity of search queries. Specifically, we collect Google Trends data for the query "bitcoin". The Google indicator tracks well the evolution of the bitcoin price. The third measure is equal to the capital raised in ICOs, expressed in dollars and as a percentage of the total bitcoin market capitalization. Most of the ICO activity is concentrated in 2017 when around 3.1 billion U.S. dollars were raised. The total capital raised, in the sample, is approximately equal to 8 billion U.S. dollars (see, for example, Lyandres et al. (forthcoming)). The fourth factor is based on the exchange bitcoin wallets, measured as a fraction of the total supply of bitcoin. It captures the net inflows of bitcoin to a particular market. Since the beginning of the sample, wallets have been growing and reached a first peak, of around 10 percent of bitcoin supply, in April 2018; dropped to less than 5 percent in March 2019; and then increased again to reach a value of around 15 percent of bitcoin supply in September 2019. Unfortunately, the data do not cover all the exchanges in our sample; in particular, we do not have data for the Korean exchanges. Although wallet identities are self-reported (see, for a discussion, Foley et al. (2019); Makarov and Schoar (2021)), and exchanges started to identify their wallets only after the Mt. Gox bankruptcy, they are commonly taken as an indicator of the transparency and of the risk of investors losing access to their assets (in this respect, wallets also measure a counterparty risk). Figure 10 summarizes the properties of these four measures.

4.2 Time-series Regressions

We regress the cross-sectional standard deviation of the log discounts across all exchange and pairs (see also Figure 5) on the set of factors described in the previous section which we group in four groups: liquidity; counter-party risk; blockchain; and cryptocurrency. We consider regressions at the weekly frequency. Specifically, we estimate the following model:

$$\mathbf{Std}_{m,j}[d_{m,j,t}] = \alpha + \beta X_t + \epsilon_t + \gamma Z_{t-1}, \tag{9}$$

Figure 10: Cryptocurrency Factors



Notes: This figure plots the bitcoin price; the average Google Trend index for the search query of the word "bitcoin" across different regions; and the value of capital raised in ICOs, as a percentage of the total bitcoin market capitalization. Data are daily from Google Trends, icobench.com, and Cryptocompare.com for the period 1/1/2014–1/1/2020.

where X is a matrix containing the explanatory factors and Z is a matrix containing further lagged explanatory variables. Table 6 presents our estimates for the sample of fiat pairs (we present the results for the crypto pairs in Table A9 in the Online Appendix). We first consider each group of factors separately (columns 1 to 4), and then in column 5 we include only the factors, from all groups, which were found to have a significant effect, and finally in column 6 we further include the lag of the cross-sectional standard deviation of the log discounts. We summarize our results as follows. First, when we consider the factors from each group separately, we find that the model explain at most 8% of the variation in the dependent variable. Second, considering liquidity factors, we find that a higher mean bid-ask spread, which likely reduces the arbitrage profits, is associated with a larger cross-section standard deviation in discounts. Third, considering counterparty factors, we find that higher values for bankruptcy shocks and DDoS are associated with a lower cross-section standard deviation in discounts, probably because they are associated with a lower number of available pairs to trade. Fourth, considering blockchain factors, we find that a higher congestion of the bitcoin blockchain, proxied by a higher volume of transactions, is associated with a larger cross-section standard deviation in discounts, possibly because it could slow the activity of arbitrageurs increasing market segmentation. Mining activity and the hash rate are also negatively related to the cross-sectional standard deviation in discounts. One possible interpretation is that mining fees and the hash rate (a measure of the available computational power) provide liquidity to local markets, reducing the dispersion in bitcoin discounts. Fifth, we find that cryptocurrency factors, and specifically higher wallet balances or investor attention are significantly related to larger cross-sectional standard deviation in discounts. We further investigate the possible explanations of these effects in the panel regressions. Sixth, in column 5 we consider a specification which include only the relevant factors for each of the four groups of factors and obtain a larger R-squared of 15%. Finally, in the last column, we further include the lagged of the dependent variable which contributes to increase the R-squared to almost 80%. Mining fees is the only factor with a significant contribution after we account for the persistence of the dependent variable. In Table A9 we present the estimates for time series regression using a sample of crypto pairs, which confirm the importance of the contribution of mining fees.

4.3 Panel Regressions

The decomposition of bitcoin discounts, over time and across exchanges and currency pairs, presented in Section 3, shows that, although the largest component is the variability of bitcoin discounts across locations, its relative contribution is time-varying, and somewhat declining over time. The results of the time-series regressions, presented above, show that, after accounting for its persistence, the cross-sectional variance in bitcoin log discounts is mostly related to variability in the mining activity.

In this section, we present the results of panel regressions to explore the time-varying determinants of discounts across locations. We assign each exchange to a geographical location g and compute the mean location-specific price relative to the aggregate bitcoin price (ψ_g). We

Table 6: Time-series Regressions: Fiat Pairs

	(1)	(2)	(3)	(4)	(5)	(6)
Liquidity						
Bid-ask	0.208**				0.118	0.053
	(1.969)				(1.244)	(0.784)
Volume	−2.082					
	(−0.198)					
Counterparty risk						
Bankruptcy shocks		−0.030***				
		(−3.173)				
DDoS		−0.120***			−0.131***	−0.017
		(−5.382)			(−5.996)	(−1.230)
BTC stolen		0.009				
		(0.086)				
Regulatory shocks		−0.096				
		(−0.284)				
Blockchain						
Time-to-transact			−0.096			
			(−0.609)			
Transactions			0.270***		0.312**	0.042
			(2.688)		(2.181)	(0.574)
Mining fees			−0.113		−0.106	−0.110*
			(−1.031)		(−0.975)	(−1.822)
Hash-rate			−0.449**			
			(−2.378)			
Cryptocurrency						
BTC return				0.003		
				(0.177)		
Investor attention				0.002**	−0.009	0.032
				(2.189)	(−0.065)	(0.475)
Wallets				0.007***	0.734***	0.109
				(4.198)	(4.214)	(1.531)
ICOs				0.002		
				(0.022)		
Persistence						0.862***
						(28.693)
R^2 (%)	0.760	4.984	4.350	8.077	15.258	78.885
Obs	369	369	369	368	369	368

Notes: The table reports the results of time-series regressions of the cross-sectional variance of log discounts. We group the covariates in four groups: liquidity risks (bid/ask spread and volume of transactions); counter-party risks (bankruptcy dummy, regulatory shock dummy, amounts of bitcoin stolen in hacks, and change in exchange wallets); blockchain factors (time-to-transact on the blockchain, the total value of transactions on the blockchain, mining activity and change in hash rate); and cryptocurrency factors (change in bitcoin price, bitcoin momentum, change in the Google Trend index, and capital raised in ICOs). Bitcoin momentum is the one-period lagged bitcoin return. The first two columns report results of regressions on daily frequency data. The third and fourth columns report results of regressions on weekly frequency data. We also report robust standard errors in parenthesis and the adjusted R-squared. Data are daily from cryptocompare.com, walletexplorer.com, Auer and Claessens (2018), Cointelegram, https://www.hackmageddon.com/, bitinfocharts.com, blockchain and coinmetrics.io, and Thomson Reuters for the period 1/1/2014–1/1/2020.

further assign each geographical location to one of two groups. We label the first group “open”, corresponding to locations with relatively low capital controls and restrictions to international flows. We label the second group “closed”, corresponding to locations with relatively higher level of capital controls. Our choice is motivated by the observation that the existence of frictions that could limit the arbitrage activity between different pairs is a necessary, but not sufficient, condition for the existence of bitcoin discounts. In order to assign locations into one of these two groups, we use the last available values of the K-control index of tightness of capital controls constructed by Fernández et al. (2016), also used by Makarov and Schoar (2020). Specifically, we assign locations with an index value below (above) the 0.15 to the closed (open) group[14].

We estimate the following model at the weekly frequency:

$$\psi_{g,t+k} = \alpha_j + \gamma_t + \mathbb{1}_o \beta_o X_{g,t} + (1 - \mathbb{1}_o)\beta_c X_{g,t} + \epsilon_{g,t+k}, \tag{10}$$

where $k = 0, \ldots, K$; X is a matrix containing a set of explanatory variables; $\mathbb{1}_o$ an indicator variable corresponding to the group of “open” locations; and β_o, β_c the vectors of regression coefficients. All the explanatory variables are at the location level. With the exception of *DDoS*, they are constructed by first computing the weekly deviation of each explanatory variable from its past eight-week average and then standardizing the resulting time-series. *DDoS* is the fraction of exchanges not available to investors in a given location.

Table 7 presents the estimation results for four specification of (10) based on the sample of fiat-to-crypto pairs and a contemporaneous relationship (i.e., $k = 0$ in (10)). Table 8 presents the results for predictive regressions (i.e., $k = 1, \ldots, 4$) to address concerns about the possible endogeneity of the explanatory variables.

We start discussing the results in Table 7 and thus we focus on conditional correlations. The first specification includes only the set of crypto determinants, which contribute to explain 2.6%

[14] The geographical locations in the “open” group are: Australia, Canada, European Union, the UK, Japan, the US, Singapore and Hong Kong. In the “closed’ group are: China, East Asia, South Korea, Latin America, Russia, Turkey and South Africa. For geographical locations including more than one country, like the EU, we consider the cross-country average of the K-control index.

of the variation in discounts across locations and time. The second specification further includes the non-crypto financial determinants which increase the regression R-squared by only 0.5% (i.e., to 3.1%). The third specification accounts for the persistence of bitcoin discounts by also including the one-week lagged values of the dependent variables which increases the overall explanatory power by a factor of 10 (i.e., to 26.5%). Finally, the fourth specification also accounts for the common—across locations—time-variation in bitcoin discounts by including a week fixed-effect which brings the regression R-squared to 45%. Our preferred specification is the third, which includes both crypto and non-crypto determinants as well as the lag of the dependent variable. This specification explains a large fraction of the variation in the data and can be directly extended to consider predictive regressions.

In closed locations the activity of cross-locations arbitrageurs is more difficult because of the tighter capital controls. As a result, we expect discounts to be more persistent and more sensitive to local shocks in these locations. In fact, limits to arbitrage are a necessary but not sufficient condition for discounts to exist. For example, variability across locations of investor attention, which is associated with demand of bitcoin, should be related to the variability of discounts across locations.

Consistent with this intuition, we find that the coefficient associated with investor attention is positive and significant for closed locations in all specifications. A one standard deviation increase in the investor attention measure is associated with an increase in the bitcoin discount of 8.2% (model 3). The effect of an increase in investor attention on bitcoin discount goes beyond the contemporaneous relationship and remains positive and statistically significant up to a lag of 3 weeks (see Table 8). In fact, we find that a one standard deviation increase in the investor attention leads to a 13% in bitcoin discounts after three weeks. We also find that wallet balances are positively and significantly related to bitcoin discounts in closed locations, but only contemporaneously and at the 10% confidence level. A one standard deviation increase in the bitcoin balance in the wallets in a given location is associated with an increase in bitcoin discount of 4.3%. This suggests that the existence of bitcoin discounts in one location drives-in arbitrageurs' crypto capital. In contrast, both investor attention and wallets are not significantly related with bitcoin discounts in open locations. We further find that mining fees are positively and significantly related with bitcoin discounts in open locations only contemporaneously: relatively higher bitcoin prices are an incentive for miners and increase their activity. In open locations, a one standard deviation in the number of bitcoins mined is associated with a reduction in bitcoin discount by 8.8%. In contrast, we find that mining fees are significantly and negatively related to bitcoin discounts in closed locations with lags of one to three weeks. For example, in closed locations a one standard deviation increase in the number of bitcoins mined in the past week is associated with a reduction in bitcoin discounts by 1.7% (see column 4 in Table 8). The lagged effect in closed locations can be interpreted in terms of an increase in location-specific bitcoin supply proxied by the mining fees. Finally, we find that the bid-ask spread, a standard measure of liquidity, is associated with a similar higher bitcoin discounts in both open and closed locations. A one standard deviation increase in the bid-ask spread is associated with an approximate increase in the bitcoin discount of 8% in both closed and open locations. A higher bid-ask spread reduces the profits of arbitrageurs and reduces their activity preventing bitcoin discounts to disappear.

Moving to the non-crypto determinants, we find that a higher domestic T-bill rate and a lower inflation are significantly associated with larger bitcoin discounts. The difference between the T-bill rate and the inflation is a measure of real funding liquidity and, as a result, a measure of the cost of raising capital for arbitrageurs. For example, a one standard deviation increase in the domestic T-bill rate is associated, in closed locations, with a 4.1% increase in the bitcoin discount. Further, we find that the domestic equity market return is significantly and negatively associated with bitcoin discounts in open locations. One possible interpretation is that periods of low domestic equity returns are associated with lower flows of capital in that specific location (for example, because of an increase in risk or risk aversion) and less arbitrage activity. While the effect of the domestic T-bill rate persists up to a 3-week lag, the effect of both domestic inflation and equity return is limited to the contemporaneous regression specification.

We further show that bitcoin discounts are persistent also at the location level. In fact, the one-week lag of the location-specific bitcoin discount is positive and significant and explains a large fraction of the total variation even controlling for the time-week-effect. The coefficients are equal to 0.43 for the open locations and to 0.56 for the closed locations. The higher coefficient for the closed locations is consistent with the fact that in these locations the activity of arbitrageurs is reduced because of the restrictions to capital flows.

Table 8 presents the estimates of Equation (10) when $k = 1, \ldots, 4$. We already commented the predictive effects of investor attention, wallets, mining and non-crypto determinants in the paragraph above. We further highlight that coefficient associated with the lagged bitcoin discounts are significant up to a lag of four week for both the open and closed locations (the only exception is the coefficient for closed location with a lag of three weeks). Interestingly, the coefficients are first positive, and then negative, indicating a reversal effect in the bitcoin discounts.

Table 9 and 10 presents the estimation results of Equation (10) for the crypto-to-crypto pairs (e.g., bitcoin-to-ethereum). The overall explanatory power is smaller than for fiat-to-crypto pairs. For example, for model 3 we obtain a R-squared of 7.36% for crypto pairs and of 26.66% for fiat pairs. Note that in the regressions for the crypto pairs we additionally include the fiat-to-crypto discount.

First, we first notice that, on average, crypto discounts tend to be lower than discounts for fiat pairs and, furthermore, that crypto discounts in exchanges located in closed economies tend to be larger than those in exchanges located in open economies and less persistent (see Table A8 in the Online Appendix). Intuitively, restrictions to capital flows and limits to arbitrage are less likely to be binding or relevant for crypto pairs and, then, the effect of local demand and local supply is smaller.

Second, looking at the estimates from the panel regressions, we find that local demand shocks, proxied by investor attention, do not have a significant effect on discounts. In contrast, local supply, proxied by mining, is associated with a significant effect. Third, we include in the regression specification also the contemporaneous and lagged values of the average fiat discount. In this case, the relationship is negative. We observe a significant effect only contemporaneously, while

Table 7: Panel Estimates for the Fiat Pairs

k-control	open	closed	open	closed	open	closed	open	closed
model	(1)		(2)		(3)		(4)	
panel A: crypto determinants								
Investor attention	−0.074	0.067**	−0.076	0.070**	−0.035	0.082***	0.015	0.119***
	(-1.039)	(2.245)	(-1.071)	(2.348)	(-0.863)	(3.935)	(0.248)	(3.221)
Mining fees	−0.143***	0.016	−0.144***	0.007	−0.088***	−0.008	−0.078***	0.006
	(-5.295)	(1.019)	(-5.026)	(0.507)	(-3.166)	(-1.358)	(-2.604)	(0.443)
Bid-ask	0.089**	0.122**	0.086**	0.122**	0.079***	0.089**	0.045	0.087**
	(2.162)	(2.277)	(2.111)	(2.116)	(2.601)	(2.091)	(1.406)	(2.038)
Wallets	0.001	0.112***	−0.000	0.116***	−0.009	0.043*	−0.014	0.056
	(0.020)	(3.198)	(-0.004)	(3.352)	(-0.349)	(1.929)	(-0.727)	(1.628)
Stolen	−0.047*	−0.006	−0.046*	−0.010	−0.048*	0.001	−0.040	0.013
	(-1.906)	(-0.157)	(-1.829)	(-0.250)	(-1.668)	(0.056)	(-1.190)	(0.475)
DDoS	−0.134	0.254	−0.126	0.332	−0.139	0.288	−0.127	0.303
	(-0.907)	(0.389)	(-0.834)	(0.535)	(-1.630)	(0.513)	(-1.375)	(0.581)
panel B: non-crypto financial determinants								
Equity market			−0.054***	−0.002	−0.055***	−0.005	−0.045*	0.024
			(-3.343)	(-0.080)	(-2.589)	(-0.176)	(-1.801)	(0.887)
T-bill			0.021	0.044*	0.023*	0.041**	0.026*	0.014
			(0.929)	(1.761)	(1.654)	(2.573)	(1.931)	(0.936)
Inflation			0.006	−0.077***	−0.002	−0.034*	0.002	−0.011
			(0.163)	(-2.822)	(-0.091)	(-1.799)	(0.093)	(-0.508)
panel C: lagged determinants								
Discount lagged					0.456***	0.524***	0.434***	0.561***
					(11.226)	(9.878)	(8.976)	(9.554)
Week-effect	No		No		No		Yes	
R^2 (%)	2.641		3.158		26.660		45.482	
Obs	4270		4270		4254		4254	

Notes: The table presents the estimates for the panel regressions of the log discounts for the fiat-to-crypto pairs as in model Equation (10). We group the determinants which we consider in three groups: the crypto determinants (panel A), the non-crypto financial determinants (panel B) and lagged determinants (panel C). For each model specification, we report the coefficients associated with the dummy for *open* and *closed* economy location in two columns. The first model (1) includes only the crypto determinants. The second model (2) additionally includes the non-crypto financial determinants. The third model (3) further includes the one-week lagged values of the dependent variable. Finally, the fourth model (4) also includes a week-fixed-effect. In parenthesis we report robust standard errors and we denote with ***,**,* significance, respectively, at the 1%, 5% and 10% level. All regressors, with the exception of *Stolen*, are constructed by first computing the weekly deviation of each determinant from its past eight-week average and then standardizing the resulting time-series. Data are weekly from Cryptocompare.com, Walletexplorer.com, Auer and Claessens (2018), Cointelegram, Hackmageddon.com, Bitinfocharts.com, Blockchain.com and Coinmetrics.io, Fernández et al. (2016), and Thomson Reuters for the period 1/1/2014–1/1/2020.

Table 8: Panel Estimates for the Fiat Pairs (all lags)

	+0		+1		+2		+3		+4	
k-control	open	closed	open	closed	open	closed	open	closed	open	closed
panel A: crypto determinants										
Investor attention	-0.035	0.082^{***}	-0.030	0.044^{**}	-0.064^{*}	0.086^{***}	-0.057	0.134^{**}	-0.046	0.107
	(-0.863)	(3.935)	(-1.413)	(2.440)	(-1.817)	(3.099)	(-1.092)	(2.137)	(-1.010)	(1.519)
Mining fees	-0.088^{***}	-0.008	-0.044	-0.017^{***}	-0.034	-0.029^{***}	-0.027	-0.018^{**}	0.023	0.004
	(-3.166)	(-1.358)	(-1.583)	(-4.137)	(-1.045)	(-4.377)	(-1.182)	(-2.318)	(0.817)	(0.420)
Bid-ask	0.079^{***}	0.089^{**}	0.050^{**}	-0.004	0.069^{**}	-0.016	0.011	-0.016	-0.061^{***}	-0.015
	(2.601)	(2.091)	(1.993)	(-0.236)	(2.501)	(-0.432)	(0.375)	(-0.565)	(-2.603)	(-0.562)
Wallets	-0.009	0.043^{*}	-0.006	0.002	-0.009	-0.020	0.002	-0.041	0.031	-0.028
	(-0.349)	(1.929)	(-0.433)	(0.084)	(-0.520)	(-0.367)	(0.101)	(-0.547)	(1.391)	(-0.374)
Persistence	0.456^{***}	0.524^{***}	0.454^{***}	0.531^{***}	0.123^{***}	0.238^{***}	-0.038^{**}	0.002	-0.148^{***}	-0.171^{***}
	(-0.349)	(1.929)	(-0.433)	(0.084)	(-0.520)	(-0.367)	(0.101)	(-0.547)	(1.391)	(-0.374)
panel B: non-crypto financial determinants										
Equity market	-0.055^{***}	-0.005	-0.025	0.001	-0.020^{*}	-0.001	-0.011	0.004	-0.015	-0.010
	(-2.589)	(-0.176)	(-1.373)	(0.095)	(-1.815)	(-0.086)	(-0.595)	(0.237)	(-0.772)	(-0.444)
T-bill	0.023^{*}	0.041^{**}	0.027^{***}	0.044^{*}	0.028^{**}	0.061^{*}	0.009	0.074^{***}	0.005	0.042^{*}
	(1.654)	(2.573)	(2.590)	(1.826)	(2.096)	(1.763)	(0.331)	(2.802)	(0.146)	(1.676)
Inflation	-0.002	-0.034^{*}	-0.007	-0.018	-0.026	0.011	-0.046	0.029^{*}	-0.006	0.021
	(-0.091)	(-1.799)	(-0.450)	(-0.956)	(-1.155)	(0.575)	(-1.474)	(1.757)	(-0.213)	(0.845)
R^2 (%)	26.660		25.432		4.676		1.384		3.620	

Notes: The table presents the estimates for the panel regressions of the log discounts for the fiat-to-crypto pairs as in model Equation (10). We group the determinants which we consider in three groups: the crypto determinants (panel A), the non-crypto financial determinants (panel B) and lagged determinants (panel C). For each model specification, we report the coefficients associated with the dummy for *open* and *closed* economy location in two columns. The first model (1) includes only the crypto determinants. The second model (2) additionally includes the non-crypto financial determinants. The third model (3) further includes the one-week lagged values of the dependent variable. Finally, the fourth model (4) also includes a week-fixed-effect. In parenthesis we report robust standard errors and we denote with ***,**,* significance, respectively, at the 1%, 5% and 10% level. All regressors, with the exception of *Stolen*, are constructed by first computing the weekly deviation of each determinant from its past eight-week average and then standardizing the resulting time-series. Data are weekly from Cryptocompare.com, Walletexplorer.com, Auer and Claessens (2018), Cointelegram, Hackmageddon.com, Bitinfocharts.com, Blockchain.com and Coinmetrics.io, Fernández et al. (2016), and Thomson Reuters for the period 1/1/2014–1/1/2020.

the effect is not significant at all lags. Recall that with no limits to arbitrage we should not observe discounts in the first place. The fact that we do observe non-zero, although small, discounts is explained by the fact that arbitrageurs might require some time to execute their trade and by the fact that these trades are risky (see, for example, Borri and Shakhnov (2022)). Moreover, if the bid-ask spread is large (which we do not observe for the crypto pairs), then arbitrage trades might not be profitable for small discounts. Nevertheless, the fact that we find a significant effect only contemporaneously indicate that arbitrageurs act relatively quickly. One possible interpretation of this contemporaneous significant effect is the following. Take, as an example, the bitcoin-to-usd pair traded in an exchange located in a closed economy and assume that its price is relative high with respect to the average price. Investors in the closed economy will, then, buy another crypto pair, for the example the ethereum-to-bitcoin pair, by selling bitcoin. On the contrary, consider the same pair traded in an exchange located in an open economy. In this case we find a significant *positive* effect indicating that also the bitcoin-to-crypto is relatively expensive. One interpretation is that bitcoin-to-usd pairs tend to be more expensive at times when investors have more appetite for risk and, at these times, investors desire a higher exposure to crypto and, mostly, to bitcoin, the main cryptocurrency by market capitalization and trading volume. In this case, also the price of bitcoin-to-crypto is then bid up. Finally, we further find a significant and negative effect associated with wallets, which we interpret as a supply effect, and a negative and significant effect associated with T-bill in exchanges located in closed economies. We relate the latter effect to the fact that the T-bill rate is an opportunity cost and higher rates discourage investors into increasing their exposure to risky securities, like bitcoin.

5 Model

In this section, we lay down a simple model consistent with the documented characteristics of bitcoin price dispersion to guide the interpretation of the empirical results. The model delivers market segmentation *in the short run* and market integration *in the long run*, consistent with the

Table 9: Panel Estimates for the Crypto Pairs

k-control	open	closed	open	closed	open	closed	open	closed
model	(1)		(2)		(3)		(4)	
panel A: crypto determinants								
Investor attention	0.081^{**}	-0.118^{***}	-0.019	0.107	-0.014	0.125	0.007	0.169
	(2.537)	(-4.261)	(-0.785)	(1.023)	(-0.732)	(1.281)	(0.147)	(1.253)
Mining fees	-0.016	0.139	0.021	-0.183^{***}	0.027	-0.241^{***}	0.068^{*}	-0.239^{***}
	(-0.682)	(1.213)	(1.098)	(-11.063)	(1.245)	(-8.794)	(1.844)	(-4.966)
Wallets	-0.052^{***}	0.000	-0.006	-0.052^{**}	-0.010	-0.088^{***}	0.000	-0.087^{***}
	(-4.547)	(NaN)	(-0.365)	(-2.455)	(-0.629)	(-4.575)	(0.020)	(-6.086)
Stolen	-0.009	-0.045^{**}	-0.016	0.003	-0.010	0.000	-0.033	-0.015
	(-0.453)	(-2.078)	(-0.343)	(0.441)	(-0.187)	(0.054)	(-0.623)	(-1.346)
DDoS	-0.012	0.001	-0.056	-0.235	-0.074	-0.184	-0.076	-0.295
	(-0.237)	(0.057)	(-0.311)	(-0.928)	(-0.555)	(-0.722)	(-0.526)	(-1.365)
Discount fiat-to-crypto pairs					0.075^{**}	-0.129^{***}	0.060^{*}	-0.115^{***}
					(2.330)	(-3.321)	(1.786)	(-3.429)
panel B: non-crypto financial determinants								
Equity market			0.025^{*}	-0.014	0.021	-0.008	0.032	0.014
			(1.785)	(-0.435)	(1.348)	(-0.300)	(1.146)	(0.581)
T-bill			-0.034	-0.089^{***}	-0.023	-0.078^{**}	-0.023	-0.078^{***}
			(-1.165)	(-2.926)	(-0.980)	(-2.530)	(-0.919)	(-3.364)
Inflation			-0.023	0.021	-0.014	0.022	-0.024	0.034^{*}
			(-0.693)	(0.990)	(-0.475)	(1.126)	(-0.799)	(1.711)
panel C: lagged determinants								
Persistence					0.260^{***}	0.186^{***}	0.281^{***}	0.200^{***}
					(4.621)	(5.317)	(4.905)	(7.469)
Discount fiat-to-crypto lagged					-0.012	0.028	-0.046	0.009
					(-0.389)	(0.813)	(-1.273)	(0.237)
Week-effect	No		No		No		Yes	
R^2 (%)	1.256		0.896		7.368		19.464	
Obs	3101		3180		3094		3094	

Notes: The table presents the estimates for the panel regressions of the log discounts for the crypto-to-crypto pairs as in model Equation (10). We group the determinants which we consider in three groups: the crypto determinants (panel A), the non-crypto financial determinants (panel B) and lagged determinants (panel C). For each model specification, we report the coefficients associated with the dummy for *open* and *closed* economy location in two columns. The first model (1) includes only the crypto determinants. The second model (2) additionally includes the non-crypto financial determinants. The third model (3) further includes the one-week lagged values of the dependent variable and of the discounts for the fiat-to-crypto pairs. Finally, the fourth model (4) also includes a week-fixed-effect. In parenthesis we report robust standard errors and we denote with ***,**,* significance, respectively, at the 1%, 5% and 10% level. All regressors, with the exception of *Stolen*, are constructed by first computing the weekly deviation of each determinant from its past eight-week average and then standardizing the resulting time-series. Data are weekly from Cryptocompare.com, Walletexplorer.com, Auer and Claessens (2018), Cointelegram, Hackmageddon.com, Bitinfocharts.com, Blockchain.com and Coinmetrics.io, Fernández et al. (2016), and Thomson Reuters for the period 1/1/2014–1/1/2020.

Table 10: Panel Estimates for the Crypto Pairs (all lags)

	+0		+1		+2		+3		+4	
k-control	open	closed	open	closed	open	closed	open	closed	open	closed
panel A: crypto determinants										
Investor attention	−0.014	0.125	−0.021	0.055	−0.008	0.028	0.022	−0.001	0.013	-0.061^{**}
	(-0.732)	(1.281)	(-1.278)	(0.608)	(-0.447)	(0.291)	(0.753)	(-0.032)	(0.624)	(-2.510)
Mining fees	0.027	-0.241^{***}	0.041	-0.057^{***}	0.016	-0.189^{***}	0.004	-0.148^{***}	−0.016	-0.031^{**}
	(1.245)	(-8.794)	(1.337)	(-4.397)	(0.536)	(-17.351)	(0.200)	(-30.218)	(-1.175)	(-2.312)
Wallets	−0.010	-0.088^{***}	-0.037^{*}	-0.132^{***}	−0.021	−0.102	−0.001	−0.028	0.020	−0.034
	(-0.629)	(-4.575)	(-1.801)	(-2.835)	(-0.820)	(-1.274)	(-0.068)	(-0.493)	(1.384)	(-0.576)
panel B: non-crypto financial determinants										
Equity market	0.021	−0.008	−0.005	0.015	−0.002	0.031	0.023	−0.012	−0.008	−0.004
	(1.348)	(-0.300)	(-0.275)	(0.579)	(-0.069)	(1.323)	(1.183)	(-0.380)	(-0.356)	(-0.248)
T-bill	−0.023	-0.078^{**}	−0.002	-0.054^{**}	−0.018	-0.065^{***}	−0.024	−0.021	0.005	0.003
	(-0.980)	(-2.530)	(-0.139)	(-2.036)	(-0.944)	(-2.850)	(-1.381)	(-0.739)	(0.352)	(0.115)
Inflation	−0.014	0.022	−0.018	0.039^{*}	0.008	0.047	0.023	0.053^{**}	0.038	0.061^{**}
	(-0.475)	(1.126)	(-0.706)	(1.665)	(0.355)	(1.587)	(0.985)	(2.271)	(1.429)	(2.174)
Persistence	0.260^{***}	0.186^{***}	0.262^{***}	0.172^{***}	0.128^{***}	0.030	-0.065^{**}	-0.078^{**}	-0.185^{***}	-0.159^{***}
	(4.621)	(5.317)	(4.649)	(5.657)	(2.953)	(1.041)	(-2.519)	(-2.555)	(-7.270)	(-5.327)
Discount fiat-to-crypto pairs	0.075^{**}	-0.129^{***}	0.021	−0.073	0.019	−0.045	0.015	−0.044	−0.004	0.011
	(2.330)	(-3.321)	(0.829)	(-1.173)	(0.819)	(-1.266)	(0.476)	(-1.450)	(-0.151)	(0.341)
R^2 (%)	7.368		6.872		2.099		1.006		4.178	

Notes: The table presents the estimates for the panel regressions of the log discounts for the crypto-to-crypto pairs as in model Equation (10). We group the determinants which we consider in three groups: the crypto determinants (panel A), the non-crypto financial determinants (panel B) and lagged determinants (panel C). For each model specification, we report the coefficients associated with the dummy for *open* and *closed* economy location in two columns. The first model (1) includes only the crypto determinants. The second model (2) additionally includes the non-crypto financial determinants. The third model (3) further includes the one-week lagged values of the dependent variable and of the discounts for the fiat-to-crypto pairs. Finally, the fourth model (4) also includes a week-fixed-effect. In parenthesis we report robust standard errors and we denote with ***,**,* significance, respectively, at the 1%, 5% and 10% level. All regressors, with the exception of *Stolen*, are constructed by first computing the weekly deviation of each determinant from its past eight-week average and then standardizing the resulting time-series. Data are weekly from Cryptocompare.com, Walletexplorer.com, Auer and Claessens (2018), Cointelegram, Hackmageddon.com, Bitinfocharts.com, Blockchain.com and Coinmetrics.io, Fernández et al. (2016), and Thomson Reuters for the period 1/1/2014–1/1/2020.

evidence of persistent, but mean-reverting, discounts. Specifically, while local supply and demand determine the price for each sub-market in the short run, bitcoin prices are equalized across markets in the long run. A calibrated version of the model matches the observed bitcoin price dispersion and persistence in closed and open economies.

5.1 Framework

We consider a repeated static model with heterogeneous investors, a market-maker trading mechanism, and J sub-markets, which correspond to the exchange locations of our empirical analysis. The model relays on three assumptions. First, investors have heterogeneous beliefs about bitcoin values. This assumption generates a reason for trade like, for example, in Harrison and Kreps (1978). Second, markets are segmented, at least in the short run. This assumption generates price dispersion, like in Grossman and Miller (1988). Third, a market-maker, like a large institutional investor or arbitrageur, visits each sub-market infrequently à la Calvo (1983) and eventually eliminates price differences across markets. This assumption guarantees that markets are integrated in the long run. Retail investors (noise traders) trade bitcoin with each other and, when present, with the market-maker. The interaction of noise traders and the market maker is related to the classic framework of Kyle (1985).

We start with one single sub-market and focus on retail investors first. For tractability, we assume that each retail investor can hold at most one of the fixed number of available bitcoins N, and cannot hold fractions of one bitcoin (i.e., she can hold either 0 or 1 bitcoin). Therefore, each period starts with N retail investors already holding bitcoins, and an even larger number of potential investors $M \geq N$. At the beginning of each period, after observing N and M, all investors draw beliefs about the fundamental value of one bitcoin from a distribution with density $f(P)$. In the absence of the market-maker, the bitcoin price on each sub-market must equate its demand to its supply:

$$N\Phi(P) = (M - N)\left(1 - \Phi(P)\right), \tag{11}$$

where N denotes the supply of bitcoins in the sub-market, M the number of prospective buyers, and $\Phi(P)$ is the cumulative distribution function evaluated at P.[15] The left-hand side of Equation (11) represents the supply of bitcoin, which equals the number of potential sellers (N) multiplied by the fraction of sellers who are willing to sell at a price P. The right-hand side of Equation (11) represents the demand of bitcoin, which equals the number of potential buyers ($M - N$) multiplied by the fraction of buyers who are willing to buy at a price P. It is convenient to assume that beliefs are distributed independently across investors and time with exponential distribution $P \sim EXP(\lambda)$, with $\lambda > 0$. In this case, the sub-market bitcoin price is:

$$P = (1/\lambda)\,[\log(M) - \log(N)] \tag{12}$$

Intuitively, according to Equation (12), the equilibrium bitcoin price is higher the higher the dispersion in beliefs about the fundamental value (i.e., for lower λ) as in Pástor and Veronesi (2006); the larger the number of possible investors (i.e., for greater M); the lower the supply of bitcoin (i.e., for smaller N) as in Cochrane (2002).

We introduce dynamics into the model by assuming that, on each submarket g, the number of potential buyers (i.e., the local demand), as well as the number of bitcoin (i.e., the local supply), change over time according to the following autoregressive stochastic processes:

$$\log(M_{g,t+1}) \equiv m_{g,t+1} = \rho_{m,g} m_{g,t} + (1 - \rho_{m,g})\bar{m} + \sigma_{m,g}\epsilon_{m,g,t+1} \tag{13}$$

$$\log(N_{g,t+1}) \equiv n_{g,t+1} = \rho_{n,g} n_{g,t} + (1 - \rho_{n,g})\bar{n} + \sigma_{n,g}\epsilon_{n,g,t+1} \tag{14}$$

where $0 < \rho_{i,g} < 1$, $\epsilon_{i,g,t+1}$, are i.i.d. $N(0, 1)$, and $i = m, n$. We further impose that $\bar{m} = 1 + \bar{n}$. The last assumption guarantees that the unconditional mean bitcoin price in each submarket is the same and equal to $1/\lambda$.

If markets are perfectly segmented, for example, because of transaction costs or restrictions to capital flows, then the equilibrium bitcoin prices in each sub-market g and period t depend on

[15] For the price to be non-negative, the condition $M \geq N$ must be satisfied.

the local demand and supply:

$$P_{g,t} = \frac{1}{\lambda}\left[m_{g,t} - n_{g,t}\right] \quad \text{with} \quad g = 1, \ldots, N_g$$

In contrast, if all investors can trade across all sub-markets, then there is a unique bitcoin price in all markets:

$$\bar{P} = \frac{1}{\lambda}\left[\sum_{g=1}^{N_g}\left(m_{g,t} - n_{g,t}\right)\right]$$

Note that, if the number of sub-markets N_g is sufficiently large, then the average price across sub-markets is always equal to the unconditional mean investors' valuation $\bar{P} = E(P) = 1/\lambda$, because $m_g - n_g$ are independently distributed with mean one by assumption.

The role of the market-maker is to bridge the short-run segmentation and the long-run integration by slowly adjusting the excess demand of bitcoins on all sub-markets. In fact, the market-maker buys or sells bitcoin at the market price $\bar{P} = 1/\lambda$, i.e., the average price across all sub-markets, but visits each sub-market infrequently with a location-specific probability $\theta_{g,t}$. At each time t, the equilibrium bitcoin price, in sub-market g, is either $\bar{P}$ or the segmented market price $P_{g,t}$ that equates the local demand and supply:

$$\tilde{P}_{g,t} = \begin{cases} \theta_{g,t} \text{ (market maker is present)} & \bar{P} \\ 1 - \theta_{g,t} \text{ (otherwise)} & P_{g,t} \end{cases}$$

5.2 Simulation

Following our empirical analysis, we define the bitcoin discounts, in the model, as the relative price in each sub-market g with respect to the mean market price:

$$D_{g,t} = \frac{\tilde{P}_{g,t}}{\bar{P}} - 1 = (1 - \theta_t)\left[m_{g,t} - n_{g,t} - 1\right]$$

In line with the results of our panel regressions, we associate local supply with mining activities as they tend to reduce location-specific bitcoin discounts in closed economies (see Table 7); and local demand with the Google Trend index, which we found to increase location-specific bitcoin discounts for the fiat pairs in closed economies (see Table 7). We assume that local demand (m_g) is a linear function of the Google Trend index for the query "bitcoin" in different locations (with scaling parameter a_m); local supply (n_g) is a linear function of the local mining activities (fees) as a fraction of total bitcoin supply (with scaling parameter a_n), and estimate the parameters of two independent autoregressive processes for each location, as specified by Equations (13)-(14). We calibrate the values of the two scaling parameters (a_m, a_n), and the time-varying frequency with which the market-marker shows up in each sub-market ($\theta_{g,t}$), by matching three moments of the bitcoin price dispersion we observe in the data and checking that $m_g \geq n_g$.

The parameter $\theta_{g,t}$ measures the probability with which the market marker visits each sub-market and can, therefore, be interpreted as a measure of the limits to arbitrage across sub-markets and over time. Given the current blockchain technology, investors cannot transfer instantaneously across exchanges. In fact, the blockchain latency implies a waiting time of at least one hour before transferring balances across exchanges, which increases with the congestion of the network. This motivates the assumption of infrequent transactions of the market-maker in the model. Although investors can buy and short-sell bitcoin, and keep balances, in different exchanges, short-selling in our sample is not available on many exchanges and the high and time-varying volatility of cryptocurrency exposes investors to substantial inventory risks. The lower the $\theta_{g,t}$ the more severe are the limits to arbitrage, and when $\theta_{g,t} = 1$ for all t, then the market-maker is always present in sub-market g and discounts are always zero.

Motivated by the panel regressions, in which we split our sample on the basis of the index of capital controls into closed and open economies, in what follows we use three samples for the calibration of the parameters $\theta_{g,t}$, a_m, a_n: the sample containing all countries; the sample containing the closed economies and the sample containing the open economies. We expect $\theta_{g,t}$ to be lower in closed economies which are associated with tighter capital controls. Specifically, we consider the following targets: the time-varying cross-sectional variation in discounts, across all locations, and the variance and autocovariance, with respect to the time dimension, of the mean discounts across locations:[16]

$$\mathbf{Var}_t D_t^g = \left[\mathbf{Var}_t\left[\theta_t\right] + (1-\mathbf{E}_t[\theta_t])^2\right]\left[\mathbf{Var}_t[m_{g,t}] + \mathbf{Var}_t[n_{g,t}]\right] \tag{15}$$

$$\mathbf{Cov}_t\left[D_t^g, D_{t-1}^g\right] = \left[\mathbf{Cov}_t\left[\theta_t, \theta_{t-1}\right] + (1-\mathbf{E}_t[\theta_t])^2\right]\left[\rho_{m,g}\mathbf{Var}_t[m_{g,t}] + \rho_{n,g}\mathbf{Var}_t[n_{g,t}]\right] \tag{16}$$

$$\mathbf{Var}_g D_{g,t} = \left[(1-\theta_t)\theta_t + (1-\theta_t)^2\right]\frac{1}{N_g}\sum_g\left[\mathbf{Var}_t[m_{g,t}] + \mathbf{Var}_t[n_{g,t}]\right] \tag{17}$$

where

$$\mathbf{Var}[m_g] = a_m\sigma_{m,g}^2\frac{1}{1-\rho_{m,g}^2} \quad \text{and} \quad \mathbf{Var}[n_g] = a_n\sigma_{n,g}^2\frac{1}{1-\rho_{n,g}^2}$$

Table 11: Average Statistics of Price Distribution Simulated Data

Simulated data	Standard deviation	Min	Max	90-10 ratio	50-10 ratio	90-50 ratio	Skewness	Kurtosis	AR(1)
Closed economies	0.031	0.844	1.146	1.077	1.039	1.037	0.100	5.726	0.695
All economies	0.028	0.834	1.160	1.057	1.030	1.026	-0.147	8.343	0.566
Open economies	0.022	0.809	1.158	1.007	1.007	1.000	-0.411	17.178	0.330

Notes: The table reports average statistics for simulated bitcoin gross discounts, defined as $(1 + D_{m,j,t})$ for three versions of the model calibrated, respectively, on a sample of closed economies, on a sample of all economies, and on a sample of open economies.

Table 11 presents the average statistics of the price distribution of simulated data for three versions of the model calibrated, respectively, on a sample of closed economies, on a sample of all economies, and on a sample of open economies. While the standard deviation and autocovariance are calibrated moments, the skewness and kurtosis of the simulated data are endogenously produced by the model. If we look at the simulated data for all economies, discounts approximately range from -20 to 20 percent and are roughly symmetric. Finally, the kurtosis is large and equal to 8.3, indicating tails that are ticker than in the case of a normal distribution. The latter is a direct consequence of the fact that the market-maker shows up in each sub-market only infrequently. If we compare the simulated data for closed and open economies, we observe that discounts are

[16] The empirical targets for the moments in Equations (15)-(17) correspond to the averages across the locations in the three sub-samples. All targets are based on the sample with all bitcoin-to-fiat and bitcoin-to-crypto pairs.

more volatile (3.1% vs. 2.2%) and persistent (0.695 vs. 0.330) in closed economies, consistent with the empirical evidence documented in Section 4.3. The latter results depends crucially on the fact that the calibrated value of the parameter $\theta_{g,t}$ is lower in closed economies.

6 Conclusions

This paper studies the distribution of bitcoin prices over time, and across markets and currencies, by considering 135 exchanges around the globe, where investors can trade bitcoin for different fiat and cryptocurrency. While the typical price distribution is roughly symmetric for bitcoin-to-crypto pairs, and more negatively skewed for bitcoin-to-fiat pairs, for all pairs is leptokurtic with a mean daily standard deviation of approximately 3.9 percent. We decompose the variance of discounts in three components listed by order of importance: time, location, quality and currency. The spatial dimension, captured by the different exchange locations, explains more than 50 percent of the total variability for fiat pairs. For crypto pairs, the currency component accounts for most the variability in bitcoin discounts.

In order to assess the importance of various contributing factors to the overall explanation of bitcoin price dispersion, we collect and merge data from multiple sources for both traditional factors, like liquidity, and cryptocurrency-specific factors, like counter-party risk, blockchain, and cryptocurrency factors. Since the focus of our paper is on the dispersion of bitcoin prices across markets and currencies, most of these measures are location-specific. We find that local supply and demand shocks accounts for a large fraction of the variability in bitcoin discounts, especially for fiat pairs, in closed locations. The latter denote exchanges located in countries with tighter capital controls.

We build a simple model to guide the interpretation of our empirical results. The model is based on three assumptions. First, investors have heterogeneous beliefs about bitcoin values. Second, markets are segmented, at least in the short run. Third, a market-maker infrequently and randomly visits each market and eliminates the price differences across markets. In this framework, in the short-run local discounts are driven by changes in local demand and supply, while in the long run, prices are equalized across markets. A calibrated version of the model matches the observed bitcoin price dispersion within and across open and closed locations.

As argued in Shiller (1994), the existence of a reliable market index is a necessary condition for the development of derivative markets, i.e., the market for futures and options. Derivatives are fundamental to insure efficiency of spot markets, for example allowing for short positions, and risk hedging. One of the conditions for the cryptocurrency market for derivatives and De-Fi to keep growing is the availability of a single and reliable bitcoin price index that could serve as underlying security for a multitude of contracts. The existence of bitcoin price differences across exchanges and currency pairs undermines the reliability of such an index. Understanding what causes these price differences is an important step in the direction of improving efficiency.

References

Auer, R., and S. Claessens. 2018. Regulating cryptocurrencies: assessing market reactions. *BIS Quarterly Review September* .

Augustin, P., A. Rubtsov, and D. Shin. 2020. The impact of derivatives on spot markets: Evidence from the introduction of bitcoin futures contracts .

Barbon, A., and A. Ranaldo. 2021. On The Quality Of Cryptocurrency Markets: Centralized Versus Decentralized Exchanges. *arXiv preprint arXiv:2112.07386* .

Biais, B., C. Bisiere, M. Bouvard, and C. Casamatta. 2019. The blockchain folk theorem. *The Review of Financial Studies* 32:1662–1715.

Biais, B., C. Bisiere, M. Bouvard, C. Casamatta, and A. J. Menkveld. 2023. Equilibrium bitcoin pricing. *The Journal of Finance* 78:967–1014.

Bitwise. 2019. Presentation to the U.S. Securities and Exchange Commission. Bitwise Asset Management.

Borri, N., and K. Shakhnov. 2019. Regulation spillovers across cryptocurrency markets. *Finance Research Letters* .

Borri, N., and K. Shakhnov. 2022. The cross-section of cryptocurrency returns. *The Review of Asset Pricing Studies* 12:667–705.

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

Burnside, C. 2011. The cross section of foreign currency risk premia and consumption growth risk: Comment. *American Economic Review* 101:3456–76.

Calvo, G. A. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12:383–398.

Catalini, C., and J. S. Gans. 2016. Some simple economics of the blockchain. NBER Working Paper No. 24242.

Chen, M., C. Qin, and X. Zhang. 2022. Cryptocurrency price discrepancies under uncertainty: evidence from COVID-19 and lockdown nexus. *Journal of International Money and Finance* 124:102633.

Chen, N.-f., R. Kan, and M. H. Miller. 1993. Are the Discounts on Closed-End Funds a Sentiment Index? *The Journal of Finance* 48:795–800.

Chiu, J., and T. V. Koeppl. 2019. Blockchain-Based Settlement for Asset Trading. *The Review of Financial Studies* 32:1716–1753.

Choi, K. J., A. Lehar, and R. Stauffer. 2022. Bitcoin microstructure and the kimchi premium. *Available at SSRN 3189051* .

Cochrane, J. H. 2002. Stocks as money: convenience yield and the tech-stock bubble. National Bureau of Economic Research Working Paper 8987.

Cong, L. W., Z. He, and J. Li. 2021a. Decentralized mining in centralized pools. *The Review of Financial Studies* 34:1191–1235.

Cong, L. W., Y. Li, and N. Wang. 2021b. Tokenomics: Dynamic adoption and valuation. *The Review of Financial Studies* 34:1105–1155.

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

De Jong, A., L. Rosenthal, and M. A. Van Dijk. 2009. The risk and return of arbitrage in dual-listed companies. *Review of Finance* 13:495–520.

Du, W., A. Tepper, and A. Verdelhan. 2018. Deviations from covered interest rate parity. *The Journal of Finance* 73:915–957.

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

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

Eom, Y. 2021. Kimchi premium and speculative trading in bitcoin. *Finance Research Letters* 38:101505.

Fernández, A., M. W. Klein, A. Rebucci, M. Schindler, and M. Uribe. 2016. Capital Control Measures: A New Dataset. *IMF Economic Review* pp. 548–574.

Figlewski, S., and G. P. Webb. 1993. Options, short sales, and market completeness. *The Journal of Finance* 48:761–777.

Foley, S., J. R. Karlsen, and T. J. Putniņš. 2019. Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies? *The Review of Financial Studies* 32:1798–1853.

Froot, K. A., and E. M. Dabora. 1999. How are stock prices affected by the location of trade? *Journal of Financial Economics* 53:189–216.

Gagnon, L., and G. A. Karolyi. 2010. Multi-market trading and arbitrage. *Journal of Financial Economics* 97:53–80.

Gandal, N., J. Hamrick, T. Moore, and T. Oberman. 2018. Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics* 95:86–96.

Grossman, S. J., and M. H. Miller. 1988. Liquidity and market structure. *the Journal of Finance* 43:617–633.

Harrison, J. M., and D. M. Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics* 92:323–336.

Huang, G.-Y., Y.-F. Gau, and Z.-X. Wu. 2022. Price discovery in fiat currency and cryptocurrency markets. *Finance Research Letters* 47.

Kapar, B., and J. Olmo. 2019. An analysis of price discovery between Bitcoin futures and spot markets. *Economics Letters* 174:62–64.

Kaplan, G., and G. Menzio. 2015. The Morphology Of Price Dispersion. *International Economic Review* 56:1165–1206.

Krishnamurthy, A. 2002. The bond/old-bond spread. *Journal of Financial Economics* 66:463–506.

Krückeberg, S., and P. Scholz. 2020. Decentralized Efficiency? Arbitrage in bitcoin Markets. *Financial Analysts Journal* .

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* pp. 1315–1335.

Lamont, O. A., and R. H. Thaler. 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy* 111:227–268.

Lee, C., A. Shleifer, and R. H. Thaler. 1991. Investor sentiment and the closed-end fund puzzle. *The Journal of Finance* 46:75–109.

Lee, J., and T. Oh. 2022. The Kimchi premium and bitcoin-cashing outlets. *Finance Research Letters* 50:103200.

Lehar, A., and C. A. Parlour. 2020. Miner collusion and the bitcoin protocol. *Available at SSRN 3559894* .

Lehar, A., and C. A. Parlour. 2021. Decentralized exchanges. Tech. rep., Working paper.

Levene, H. 1960. Robust tests for equality of variances. *Contributions to probability and statistics* pp. 278–292.

Liu, Y., and A. Tsyvinski. 2021. Risks and returns of cryptocurrency. *The Review of Financial Studies* 34:2689–2727.

Liu, Y., A. Tsyvinski, and X. Wu. 2022. Common risk factors in cryptocurrency. *The Journal of Finance* 77:1133–1177.

Lustig, H., and A. Verdelhan. 2007. The cross section of foreign currency risk premia and consumption growth risk. *American Economic Review* 97:89–117.

Lyandres, E., B. Palazzo, and D. Rabetti. forthcoming. ICO success and post-ICO performance. *Management Science* .

Ma, J., J. S. Gans, and R. Tourky. 2018. Market Structure in Bitcoin Mining. NBER Working Paper.

Makarov, I., and A. Schoar. 2019. Price discovery in cryptocurrency markets. *AEA Papers and Proceedings* 109:97–99.

Makarov, I., and A. Schoar. 2020. Trading and Arbitrage in Cryptocurrency Markets. *Journal of Financial Economics* 135:293–319.

Makarov, I., and A. Schoar. 2021. Blockchain analysis of the bitcoin market. Tech. rep., National Bureau of Economic Research.

Makarov, I., and A. Schoar. 2022. Cryptocurrencies and Decentralized Finance (DeFi). Tech. rep., National Bureau of Economic Research.

Moore, T., and N. Christin. 2013. Beware the middleman: Empirical analysis of Bitcoin-exchange risk. In *International Conference on Financial Cryptography and Data Security*, pp. 25–33. Springer.

Nakamoto, S. 2008. Bitcoin: A peer-to-peer electronic cash system.

Nakamura, E., J. Steinsson, R. Barro, and J. Ursúa. 2013. Crises and recoveries in an empirical model of consumption disasters. *American Economic Journal: Macroeconomics* 5:35–74.

Ok, H., J. Kim, and Y. Kim. 2023. Is the Kimchi premium a speculative bubble? *Finance Research Letters* p. 104207.

Pagnotta, E., and A. Buraschi. 2018. An equilibrium valuation of bitcoin and decentralized network assets. *Manuscript* .

Pan, J., and A. M. Poteshman. 2006. The information in option volume for future stock prices. *The Review of Financial Studies* 19:871–908.

Pástor, L., and P. Veronesi. 2006. Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81:61–100.

Shakhnov, K., and L. Zaccaria. Forthcoming. Utility Token, Network Effects, and Pricing Power. *Management Science* .

Shams, A. 2020. The Structure of Cryptocurrency Returns. *Charles A. Dice Center Working Paper* .

Shiller, R. J. 1994. *Macro markets: creating institutions for managing society's largest economic risks.* OUP Oxford.

Shiller, R. J. 2008. Derivatives markets for home prices. Tech. rep., National Bureau of Economic Research.

Vasek, M., and T. Moore. 2015. There's no free lunch, even using Bitcoin: Tracking the popularity and profits of virtual currency scams. In *International Conference on Financial Cryptography and Data Security*, pp. 44–61. Springer.

Velde, F., et al. 2013. Bitcoin: A primer. *Chicago Fed Letter* .

Wang, S. S., and L. Jiang. 2004. Location of trade, ownership restrictions, and market illiquidity: Examining Chinese A-and H-shares. *Journal of Banking & Finance* 28:1273–1297.

Yermack, D. 2013. Is Bitcoin a real currency? An economic appraisal. National Bureau of Economic Research Working Paper 19747.

Author Statement

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). He is responsible for communicating with the other authors about progress, submissions of revisions, and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author.

Authors: Kirill Shakhnov and Nicola Borri