Journal of Banking and Finance 124 (2021) 106041

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

How to measure the liquidity of cryptocurrency markets?

Alexander Brauneis[a], Roland Mestel[b,*], Ryan Riordan[c], Erik Theissen[b,d]

[a] University of Klagenfurt, Department of Finance & Accounting, Universitaetsstrasse 65-67, Klagenfurt A-9020, Austria
[b] University of Graz, Institute of Banking and Finance, Universitaetsstrasse 15/F2, Graz A-8010, Austria
[c] Smith School of Business, Queen's University, Goodes Hall, 143 Union St West, Kingston, ON K7L 3N6, Canada
[d] University of Mannheim, Finance Area, L9 1-2, Mannheim D-68161, Germany

A R T I C L E I N F O

Article history:
Received 22 July 2020
Accepted 27 December 2020
Available online 4 January 2021

JEL classification:
G12
G14

Keywords:
Cryptocurrencies
Liquidity
Capital markets

A B S T R A C T

This paper investigates the efficacy of low-frequency transactions-based liquidity measures to describe actual (high-frequency) liquidity. We show that the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators outperform other measures in describing time-series variations, irrespective of the observation frequency, trading venue, high-frequency liquidity benchmark, and cryptocurrency. Both measures perform well during high and low return, volatility and volume periods. The Kyle and Obizhaeva (2016) estimator and the Amihud (2002) illiquidity ratio outperform when estimating liquidity levels. These two estimators also reliably identify liquidity differences between trading venues. Overall, the results suggest that there is not yet a universally *best* measure but there are reasonably *good* low-frequency measures.

1. Introduction

Bitcoin and other cryptocurrencies are now firmly entrenched in the financial system. Bitcoin is becoming a widely accepted form of online payment and more than 35 million bitcoin wallets are in existence. Trading in bitcoin exceeded $930 billion USD in January 2020. Bitcoin also forms a growing part of investment portfolios and serves as the underlying for futures contracts,[1] recently exceeding $1 billion in open interest.[2] Originally designed as a decentralized digital cash system using cryptographic hash functions to secure transactions, it is poised to overtake national fiat currencies and other financial assets in terms of global importance.

Bitcoin and other cryptocurrencies are traded on numerous trading platforms around the globe. Bitcoin can be traded 24 h per day, and seven days per week for US dollars, the Euro, Japanese Yen, as well as numerous other fiat and (crypto)currencies. Monthly dollar trading volume on the New York Stock Exchange was of similar magnitude to global bitcoin trading, with $1.03 trillion USD for the NYSE during December 2019.[3]

The growing importance of bitcoin for payments and investments is dependent on an efficient transfer of bitcoin for other currencies on cryptocurrency exchanges. The number of exchanges has exploded, making it difficult for investors to select an exchange for trading and hedging. While trading has become relatively frequent in cryptocurrencies the liquidity of these markets is difficult to determine. Cryptocurrency markets lack a regulated data feed like the consolidated tape for U.S. equities. The lack of a consolidated feed, coupled with the high number of exchanges and jurisdictions makes it difficult to calculate high-frequency bid-ask spreads thereby hampering the comparison of liquidity across cryptocurrency exchanges. The bid-ask spread is an important metric when assessing an exchange in that it represents the costs of immediately buying or selling a security. Bid-ask spreads are usually calculated using high-frequency intraday data that are both expensive to purchase and time-consuming to process. We compare

* Corresponding author.
E-mail addresses: alexander.brauneis@aau.at (A. Brauneis), roland.mestel@uni-graz.at (R. Mestel), ryan.riordan@queensu.ca (R. Riordan), theissen@uni-mannheim.de (E. Theissen).
URL: http://www.aau.at/fin (A. Brauneis), https://banken-finanzierung.uni-graz.at (R. Mestel), http://ryanriordan.ca/ (R. Riordan), https://www.bwl.uni-mannheim.de/theissen/ (E. Theissen)

[1] See: https://www.cmegroup.com/trading/bitcoin-futures.html

[2] https://www.coindesk.com/bitcoin-futures-pass-1b-in-open-interest-on-bitmex-for-first-time-since-march-crash

[3] Data for bitcoin volume come from https://coinmarketcap.com/currencies/bitcoin/historical-data/ and NYSE volume data are from https://focus.world-exchanges.org/issue/january-2020/market-statistics. Both sites were accessed on February 15, 2020. We note, though, that some cryptocurrency exchanges appear to be overstating their reported volume (e.g. Hougan et al., 2019) and that ranking websites like coinmarketcap.com may be tempted to report inflated trading volume due to their revenue model which is largely dependent on crypto-exchanges (Alexander and Dakos, 2019).

https://doi.org/10.1016/j.jbankfin.2020.106041

high-frequency measures of liquidity with easy to compute low-frequency measures.

Characterising liquidity across exchanges is important for investors, traders, and hedging strategies that use cryptocurrencies (Hu et al., 2019) that can be negatively affected by the costs of illiquidity. Additionally, cryptocurrency prices are not integrated across exchanges (Makarov and Schoar, 2020) and the decision to trade on an exchange is binding as orders cannot easily be rerouted to exchanges that are more liquid or offering better prices. With little information about individual exchanges traders may have to rely on transactions data such as the daily high, low and closing prices to evaluate market quality.

We study the accuracy of liquidity measures derived from transactions data. To this end we estimate low-frequency measures derived from aggregate transactions data[4] (prices and volumes) and compare them to high-frequency measures of transaction costs and price impact calculated from order book data. Our objective is to identify the transactions-based measure that best describes actual liquidity on a cryptocurrency exchange.

Data on best bid and ask prices and order books are hard to obtain and process.[5] As such, few papers use full order book data to study the liquidity of cryptocurrency markets (Brauneis et al., 2019; Dyhrberg et al., 2018; Hautsch et al., 2018; Makarov and Schoar, 2020 and Marshall et al., 2019). A number of low-frequency measures have been developed and used to assess bond, commodity, foreign exchange and equity market liquidity (e.g. Fong et al., 2017; Goyenko et al., 2009; Karnaukh et al., 2015; Johann, Theissen, 2020; Marshall et al., 2012; Schestag et al., 2016). Cryptocurrency markets have characteristics that differ from traditional markets,[6] suggesting that liquidity formation on cryptocurrency exchanges may differ from those of other asset markets.

We use a novel and comprehensive set of continuous transactions data and order book snapshots comprising the 50 best bids and asks for two major cryptocurrencies (bitcoin and ethereum) and three large exchanges (Bitfinex, Bitstamp and Coinbase Pro) over a two-year period. First, we use these data to construct high-frequency measures of transaction costs and price impact. These measures serve as our liquidity benchmarks. In a second step, we use transactions data (prices and volumes) and calculate various liquidity proxies at lower frequencies (1 h, 1 day, and 15 days,[7] respectively). Data to compute the measures are collected at the 1-minute, 1-hour and 1-day frequency.[8] Individual low-frequency measures have been used to describe liquidity in cryptocurrency markets (e.g. Brauneis and Mestel, 2018; Dimpfl, 2017; Fink and Johann, 2014; Shi, 2018) but the relative benefits of each is not well understood. The most commonly used of these low-frequency measures are the Roll (1984) serial covariance estimator and the Amihud (2002) illiquidity ratio. We extend the analysis to include low-frequency measures based on high and low prices (Abdi and Ranaldo, 2017; Corwin and Schultz, 2012), the volatility-to-volume measure proposed by Kyle and Obizhaeva (2016) and simple volume-based measures (the transaction frequency and the USD trading volume). Our goal is to evaluate these measures' ability to describe time-series variation in liquidity as well as level differences across exchanges. This will allow us to recommend a specific liquidity proxy using easy-to-access transactions data and an easy-to-compute proxy. Using such a proxy offers enormous savings compared to high-frequency order book measures.

Our paper augments the literature on low-frequency transactions-based liquidity measures by extending the analysis to cryptocurrencies, an important and emerging asset class. We focus on results for bitcoin, the largest cryptocurrency. Results for ethereum, which are qualitatively similar in most respects, are discussed in Section 3.7, the corresponding tables and figures are in the appendix. We use the cost of a roundtrip trade in addition to quoted spreads, effective spreads and price impacts as benchmark measures. The round-trip measure provides estimates of the execution costs for large trades and is thus important for evaluating some trading strategies (Hu et al., 2019) and factor models (Liu et al., 2019).

The results suggest that the proxies that use high, low and closing prices, the Corwin and Schultz (2012) and Abdi and Ranaldo (2017)) estimators, best capture the time-series variation in cryptocurrency liquidity. These measures work for all data frequencies, exchanges (Bitfinex, Bitstamp, Coinbase Pro), benchmark measures (quoted spread, effective spread, price impact, cost of a roundtrip trade) and for both bitcoin and ethereum.[9] Average time-series correlations describe the average relationship between benchmark measures and proxies but do not capture the relationship for extreme liquidity events that may be important for investment and hedging strategies. We use quantile dependence plots to understand how well transactions-based liquidity measures capture the time-series properties of the benchmark measures across the distribution. This is an important extension since the relative performance of liquidity proxies might be different depending on the liquidity regime. We also perform several sample splits and find similar performance rankings of our liquidity proxies for high and low volume and volatility periods. Given the extreme volatility associated with bitcoin and cryptocurrency markets more generally, identifying liquidity proxies that perform well in different volatility and volume scenarios is an important contribution.

The popular Amihud (2002) illiquidity ratio does not capture the time-series variability of liquidity in the cryptocurrency markets.[10] The poor performance is driven by the relationship between volume and liquidity that is assumed to be negative in Amihud (2002) and is positive in cryptocurrency markets. The positive relation between bid-ask spreads and volume is at odds with most theoretical predictions but has recently also been documented by Bogousslavsky and Collin-Dufresne (2020) for large US stocks. [11]

Similar to Hasbrouck and Seppi (2001), we construct a composite estimator, the first principal component of the low-frequency proxies, and find that it does not generally improve on the performance of the best individual proxies.

The measures that best describe the level of the benchmark measures are the Kyle and Obizhaeva (2016) and Amihud (2002) estimators. In this application the Corwin and

[4] Below, we use the term transactions-based measure synonymously with the terms low-frequency measure and liquidity proxy.

[5] It can be downloaded in real time from the REST APIs of each cryptocurrency exchanges or it can be purchased from vendors such as Kaiko.

[6] e.g. markets are highly fragmented and weakly regulated; they are open 365 days a year and 24 h each day, they allow direct market access for all traders; trading platforms allow a direct transfer of fiat currency from and to bank accounts or credit cards, and transactions are cleared and settled by exchanges directly; margin trading and short-selling is uncommon.

[7] Providing results at the monthly frequency is infeasible because high-frequency data for cryptocurrencies are limited to 24 months.

[8] In contrast to CRSP for equity markets that provides daily prices, cryptocurrency prices are available at higher than daily frequency. For example, the site cryptodatadownload.com provides free data for many currency pairs and trading venues at the hourly frequency.

[9] Consistent with our results, Karnaukh et al. (2015) report the Corwin and Schultz (2012) measure to have the highest correlation with high-frequency bid-ask spreads in FX markets.

[10] Conceptually, the Amihud (2002) illiquidity ratio is a proxy for the price impact, not for the spread. We find that it also performs poorly in tracking the time-series variation of price impacts, a component of the effective spread.

[11] Amihud and Noh (forthcoming) present evidence of occasions where the illiquidity ratio and the inverse of volume move in opposite directions, implying that volume increases while liquidity, as measured by the illiquidity ratio, decreases.

Schultz (2012) and Abdi and Ranaldo (2017) estimators perform poorly. We find that the values obtained for these two estimators and for the Roll (1984) estimator are negatively related to the data frequency, a finding that has been documented previously for the Roll estimator (Roll, 1984; Harris, 1990) but has, to the best of our knowledge, not been documented for the high-low spread estimators.

An important application of liquidity proxies is to select an execution venue among a number of alternatives. We use the low-frequency estimators to rank trading venues according to their liquidity. We find that the Amihud (2002) illiquidity ratio and the Kyle and Obizhaeva (2016) estimator best replicate the 'true' ranking when compared to the ranking generated using high-frequency order book measures.[12]

Our findings are useful for researchers, investors, traders, trading venue operators and regulators to understand liquidity levels and dynamics on cryptocurrency exchanges with relatively easy to acquire and process aggregate price and volume data. Investors seeking the most liquid exchanges are best advised to use the Amihud (2002) illiquidity ratio or the Kyle and Obizhaeva (2016) estimator. These two measures also provide good approximations of the level of liquidity. They are the measures of choice for market participants attempting to estimate execution costs to evaluate trading strategies. In contrast, traders seeking to time the liquidity of cryptocurrency markets and enter or exit when markets are liquid should use the Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimator as they best capture the time-series variability of the quoted and effective bid-ask spread. Regulators and trading venue operators can learn from our paper about how exchanges compare across time and use aggregate measures to study the impact on liquidity of regulatory or market changes. Researchers can use our results to guide their choice of liquidity measures in empirical studies on cryptocurrency markets. Overall, our results suggest that the measure used should depend on the question being asked, as there is not (yet) a universally best measure.

The remainder of the paper is organized as follows. In Section 2 we describe our data and methodology, Section 3 presents the results, and Section 4 concludes.

2. Data and methodology

2.1. Data

We compile a high-frequency data set that covers the two-year period from 12/16/2017 00:00 UTC to 12/16/2019 00:00 UTC, a total of 730 trading days (17,520 h). Over this period we used Matlab to continuously access the public and freely accessible REST APIs of three large trading venues, Bitfinex, Bitstamp and Coinbase Pro (formerly known as GDAX). These are among the largest cryptocurrency spot trading platforms. All three venues operate an electronic central limit order book with orders being matched based on price and time priority.

The REST APIs provide live information on transactions and the current state of the order book. All public endpoints at each of these exchanges use GET requests for different types of information. We request records on 'Trades' / 'Transactions' and the 'Order book'. Depending on the venue, request parameters vary. For instance, Bitstamp only provides the full order book (with usually thousands of entries) whereas order book requests at Bitfinex and Coinbase Pro may be limited to the 50 best price levels on each side of the market.

Table 1
Number of transactions and order book snapshots for three exchanges and the two cryptocurrencies bitcoin (BTC) and ethereum (ETH), both traded against the USD. The sample period is 12/16/2017 to 12/16/2019.

	Number of transactions		Number of order books	
	BTC	ETH	BTC	ETH
Bitfinex	37,148,069	23,820,982	7,271,422	7,336,639
Bitstamp	15,310,565	5,207,030	6,913,021	6,611,326
Coinbase Pro	38,241,727	24,420,844	8,186,287	8,186,287

A potential problem associated with transactions data from cryptocurrency exchanges are fake data. A widely cited report by Hougan et al. (2019) argues that up to 95% of exchange-reported trading volume in bitcoin might not represent economically meaningful transactions or might even be plain fake. Collecting unique high-frequency trade and order book data for bitcoin the authors subject 83 cryptocurrency exchanges to several tests to identify exchanges that are likely to overstate trading volume. Only 10 exchanges passed all the tests and are characterized as "real volume" exchanges. The three trading venues that we consider in the present study all belong to the latter group.

From each trading venue we download data for two currency pairs, bitcoin versus US dollar (BTCUSD) and ethereum against US dollar (ETHUSD). The data set includes the price and the corresponding dollar trading volume for each transaction, a UNIX time stamp, a unique exchange-specific ID and a trade indicator which indicates whether a transaction was buyer-initiated or seller-initiated. Table 1 lists the total number of transactions and order book snapshots for both currencies and all three markets. A total of 90.7 (53.4) million transactions were executed for bitcoin (ethereum) during the investigation period, most of them on Coinbase Pro while Bitstamp reports least transactions.

We observe several time intervals with gaps in the data. These may be due to actually missing trading activity, technical problems (failure of the internet connection, no response from the server etc.), or exchange-specific trading halts (e.g. due to maintenance, updates or hacker attacks). We identify between 6,329 (Coinbase Pro - BTC) and 187,254 (Bitstamp - ETH) intervals without transaction data exceeding 60 s (1 min), between 2,641 and 5,920 intervals exceeding 600 s (10 min) and between 1,573 and 2,199 intervals exceeding 1,800 s (30 min). [13].

Table 2 provides descriptive statistics. Trading activity is markedly higher for BTCUSD than for ETHUSD on all three exchanges. The differences are more pronounced for the USD trading volume than for the number of transactions, implying that the average trade size is smaller for the currency pair ETHUSD. With respect to the number of transactions Coinbase Pro (Bitstamp) is the most active (least active) exchange for both currencies. However, average daily USD volume is highest on Bitfinex and lowest on Bitstamp. Concerning average USD trade size Bitstamp (Coinbase Pro) has the highest (lowest) level for the pair BTCUSD, while for ETHUSD Bitfinex (Coinbase Pro) shows the highest (lowest) level. The standard deviation of ETHUSD returns is larger than that of BTCUSD returns on all three venues. Across venues, price returns are most volatile on Bitstamp and least volatile on Coinbase Pro.[14]

Besides transactions data we retrieve order book data from the three trading platforms. Specifically, we collect the 50 best bid and best ask prices with corresponding volumes, resulting in a total of 14.6 million (13.5 million, 16.4 million) order book snap-

[12] The Corwin and Schultz (2012) estimator does very well for bitcoin but does poorly for ethereum.

[13] For Bitstamp and Coinbase Pro we have a 19 days lack of data in September 2018

[14] We note that the differences in the standard deviation of returns may reflect liquidity differences because the return standard deviation is affected by bid-ask bounce. In fact, as shown in Table 3 below, bid-ask spreads are largest on Bitstamp, a result that has also been confirmed by Brauneis et al. (2019).

Table 2
Descriptive statistics for transactions data and the quote midpoint for the pairs BTCUSD and ETHUSD. Number of transactions and dollar volume refer to daily averages, the standard deviation of price returns $\sigma(r)$ as well as the standard deviation of the quote midpoint returns $\sigma(MQ)$ refer to returns normalized to 60 s. The sample period is 12/16/2017 to 12/16/2019.

	BTCUSD				ETHUSD			
	# TX [1000]	dollar vol [mio USD]	$\sigma(r)$ [bp]	$\sigma(MQ)$ [bp]	# TX [1000]	dollar vol [mio USD]	$\sigma(r)$ [bp]	$\sigma(MQ)$ [bp]
Bitfinex	51.15	149.87	13.44	13.51	32.74	57.46	16.88	15.87
Bitstamp	21.67	66.93	15.19	12.28	7.36	11.68	26.44	14.53
Coinbase Pro	54.05	87.73	12.78	13.23	34.61	40.17	15.22	15.27

Table 3
Descriptive data for benchmark liquidity measures for the pair BTCUSD used in the empirical analysis in the empirical analysis. Values are based on our complete record of all transactions and order book snapshots and represent averages at a daily resolution. The table reports descriptive statistics for the quoted spread (QS), the effective spread (ES), the price impact (PI) and the percentage cost of a roundtrip trade (CRT). Q1 (Q3) denotes the first (third) quartile. The unit of measurement is basis points. The sample period is 12/16/2017 to 12/16/2019.

exchange		mean	std. dev.	Q1	median	Q3	num daily obs.
Bitfinex	QS	0.721	0.691	0.293	0.456	0.986	695
	ES	1.040	1.075	0.459	0.743	1.353	
	PI	0.394	0.330	0.159	0.303	0.527	
	CRT	3.833	1.927	2.351	3.649	4.677	
Bitstamp	QS	6.553	3.270	3.888	6.377	8.163	663
	ES	6.992	3.332	4.451	6.675	8.456	
	PI	0.492	0.415	0.214	0.363	0.644	
	CRT	13.23	5.765	9.601	11.88	15.70	
Coinbase Pro	QS	0.636	3.965	0.061	0.194	0.652	668
	ES	1.173	1.352	0.420	0.816	1.456	
	PI	0.395	0.491	0.120	0.252	0.485	
	CRT	3.407	4.202	2.272	2.945	3.911	

shots for Bitfinex (Bitstamp, Coinbase Pro) for the two cryptocurrencies under investigation (see Table 1). As for the transactions data we observe a considerable number of intervals without order book snapshots. There are between 13,038 (Coinbase Pro) and 84,461 (Bitfinex) intervals without data exceeding 60 s. The numbers of intervals without order book snapshots exceeding 600 s and 1,800 s are roughly equal across the three exchanges and amount to approximately 3,300 and 2,200, respectively. The standard deviations of quote midpoint returns are similar across trading venues and are generally higher for ETHUSD than for BTCUSD (see Table 2).

2.2. Measures of liquidity

The purpose of our paper is to assess and compare the accuracy of transactions-based measures of liquidity. In doing so we take the perspective of a researcher who has access to data on open, high, low and close prices and on the number of transactions and the dollar trading volume.

For our analysis we need to specify (a) the frequency at which these data are available (measured by the length of the *subintervals* i in the sequel) and (b) the frequency at which the transactions-based measures are calculated (measured by the length of the *intervals* t). Unlike for other financial markets (e.g. stock markets), price and volume data for cryptocurrencies are easily available for higher than daily frequencies. We therefore choose three distinct setups.

- Data are available at the 1-minute frequency and are used to estimate transactions-based liquidity measures at the hourly frequency.
- Data are available at the 1-hour frequency and are used to estimate liquidity measures at the daily frequency.
- Data are available at the daily level and are used to calculate liquidity measures at a 15-day frequency.

To construct the data set we use our record of all transactions and extract the open, high, low and close price as well as the number of transactions and the dollar volume at the respective frequencies of one minute, one hour and one day. These data are then used to calculate the transactions-based measures (to be described below) at the hourly, daily and 15-day frequencies, respectively. Because the three trading venues under investigation are located in different time zones we follow coinmarketcap.com and define a trading day as lasting from 00:00 UTC to 23:59 UTC. For an interval to be included in the analysis we require that data are available for at least 80% of the subintervals. Thus, when we aggregate minute-by-minute (hour-by-hour, daily) data to the hourly (daily, 15-day) frequency we require at least 48 min (19 h, 12 days) with valid data. The final data set roughly comprises 12,500 hourly intervals, 670 daily intervals and 46 15-day intervals, respectively.

Besides the transactions-based measures we need to calculate benchmark measures. To this end we use the complete record of all transactions and all order book snapshots and calculate average quoted and effective spreads, price impacts and cost of a roundtrip trade (to be defined below) at the hourly, daily and 15-day frequency.

In the sequel we first describe the high-frequency measures which we use as benchmark measures and then the transactions-based measures that we wish to evaluate.

2.2.1. High-frequency benchmark measures

- Percentage Quoted Spread (QS)
  The percentage quoted spread is the difference between the best ask price P^a and the best bid price P^b of each order book snapshot, divided by the quote midpoint $MQ = (P^b + P^a)/2$ and averaged over all observations in the interval
  $QS_t = \frac{1}{N_t}\sum_{j=1}^{N_t} \frac{P_j^a - P_j^b}{MQ_j}$.
  The subscript j denotes the j^{th} order book snapshot in interval t and N_t is the total number of order book snapshots in interval t.
- Percentage Effective Spread (ES)
  To estimate the effective bid-ask spread we combine order book snapshots with the first transaction that occurs after the snapshot.[15] The price of this transaction is denoted P^+. The average

[15] When there is more than one transaction between two order book snapshots we only use the first of these transactions. We further require the transaction to happen within 60 sec after the order book record.

percentage effective spread in interval t is then calculated as $ES_t = \frac{1}{N_t^+}\sum_{j=1}^{N_t^+}\frac{2\cdot|P_j^+ - MQ_j|}{MQ_j}$
N_t^+ is the number of order book snapshots that are followed by a transaction before the next order book snapshot is recorded.

- Percentage Price Impact (*PI*)
To estimate the price impact we use data sequences consisting of an order book snapshot, the first transaction after the snapshot and the subsequent order book snapshot. The percentage price impact is then calculated as the signed percentage change in the quote midpoint from the pre-transaction order book snapshot to the post-transaction snapshot,[16] averaged over all data sequences (as defined above) in the interval

$$PI_t = \frac{1}{N_t^+ - 1}\sum_{j=1}^{N_t^+ - 1} Q_j^+ \cdot \frac{(MQ_{j+1} - MQ_j)}{MQ_j}$$

where Q_j^+ denotes the trade indicator (+1 for a buyer-initiated trade and −1 for a seller-initiated trade) of the transaction occurring after the order book snapshot j Conrad, Wahal (2020).[17]

- Percentage cost of a roundtrip trade ($CRT(Y)$)
To assess the liquidity for larger trades we use the order book data to calculate the weighted average prices at which a buy and a sell order of a given size Y would execute. The weighted average price for executing a transaction of size Y USD given the current state of the order book is defined as $\frac{\sum_{k=1}^{K} A_k \cdot V_k}{\sum_{k=1}^{K} V_k}$ subject to $\sum_{k=1}^{K} A_k \cdot V_k = Y$ where A_k and V_k are the price and volume of the k^{th} order, respectively. Note that the K^{th} order may be subject to partial execution, depending on the outstanding dollar volume required to entirely fill the transaction volume Y. We set Y equal to the 99% quantile of the corresponding (aggregate) trade size distribution. For the currency pair BTCUSD this value is approximately equal to USD 32,100, while for ETHUSD Y roughly corresponds to USD 17,400.
To estimate the cost of a roundtrip trade of size Y, $CRT(Y)$, we calculate the weighted average prices for a market buy order and a market sell order of size Y and then express the difference between the two prices as a fraction of their midpoint. Finally, we calculate an equally-weighted average across all order book snapshot in interval t.
The $CRT(Y)$ measure is conceptually similar to the quoted bid-ask spread. However, while the quoted spread measures the transaction costs of a small trade (defined as a trade the size of which does not exceed the quoted depth), the $CRT(Y)$ measure estimates the execution costs of a trade of size Y.

Table 3 shows descriptive statistics for the benchmark measures for BTCUSD obtained from daily data (results at the hourly and 15-day frequency are virtually identical and available upon request). Overall, the percentage trading costs in the cryptocurrency markets are very low. Average quoted and effective spreads on Bitfinex and Coinbase Pro are below 1.2 bps while the cost of a roundtrip trade of size USD 32,100 on these two venues are below 4 bps.[18] The price impact on these two exchanges amounts to approximately 35% of the effective spread, implying that the suppliers of liquidity earn a small positive realized spread on average. Quoted spreads, effective spreads and the cost of a roundtrip trade are much larger and more volatile on Bitstamp than on the other two exchanges.[19] The price impact, on the other hand, is only slightly larger on Bitstamp than on the other two venues. Thus, suppliers of liquidity on Bitstamp appear to be earning significant realized spreads.

We also calculated correlations between our benchmark liquidity measures (not tabulated). Using the daily data set for the pair BTCUSD (results for other frequencies as well as for ETHUSD are similar and available upon request) we find the highest correlation between QS and CRT (exchange average: 0.89). Note that QS is a special case of CRT, measuring the cost of a roundtrip for trades not exceeding the quoted inside depth. The high correlation between QS and CRT is therefore not surprising and in line with similar results from stock markets (Irvine et al., 2000). We take it as evidence that QS is a good indicator for market liquidity not only at but also beyond the inside spread. PI has the lowest correlations with QS (exchange average: 0.49) and CRT (exchange average: 0.55). This confirms that PI captures a different dimension of market liquidity than the spread measures.

Fig. 1 shows the evolution of the hourly benchmark measures for the pair BTCUSD over time. Overall the patterns reveal strong similarities in the liquidity measures, both within and between the three exchanges. Starting from higher levels in December 2017, liquidity measures continuously decrease until the end of November 2018 where they increase sharply. Throughout 2019 liquidity is lower and more volatile than during most of 2018.

Corresponding to Table 3, Table 9 in the appendix provides descriptive statistics for the benchmark liquidity measures for the currency pair ETHUSD at the daily frequency (again, results for the other frequencies are essentially identical and are available upon request). As for BTCUSD we find the levels of our four benchmark measures to be very similar on Bitfinex and Coinbase Pro. Average quoted spreads are about twice as high as those for the pair BTCUSD. Average effective spreads for ETHUSD are below 2.2 bps on Bitfinex and Coinbase Pro, but again are higher than those for BTCUSD. Average price impacts on both trading venues amount to roughly 30% of effective spreads , again implying that the suppliers of liquidity earn a small realized spread on average.

As for BTCUSD, Bitstamp is substantially less liquid for ETHUSD than the other two exchanges.[20] The average quoted spread (effective spread) amounts to 13.17 bps (13.45 bps). Again the average

[16] From the numbers in Table 1 it follows that we observe an order book snapshot every nine seconds on average. Thus, the horizon over which we calculate the price impact is slightly less than nine seconds on average. This is in line with Conrad and Wahal (2020) who recommend to use a horizon of no more than 15 s for liquid stocks.

[17] As before, when there is more than one transaction between two order book snapshots we only use the first of these transactions. Also, we discard order book observations more than 60 sec apart. We lose one observation in each interval because the last order book snapshot in an interval is discarded as it is not followed by another snapshot in the same interval.

[18] By way of comparison: Mancini et al. (2013) report liquidity for the 9 most traded exchange rates on the EBS platform over the period January 2007 to December 2009. They find EURUSD to be the most liquid rate with a mean relative quoted spread (effective spread) of 1.05 (0.31) basis points. USDCAD is the least liquid of the analyzed pairs with respect to the quoted spread (8.27 basis points) while AUDUSD has the highest effective spread (1.38 basis points.

[19] Because of the higher execution costs on Bitstamp traders may want to avoid Bitstamp. However, there are several reasons why we may still observe significant trading activity on Bitstamp. First, most transactions are small. The median trade size on Bitstamp is 354 USD (the corresponding values for Bitfinex and Coinbase Pro are 500 and 140 USD, respectively). Assuming a quoted spread of 6.6 bps (the median quoted spread on Bitstamp), the execution costs of a median-sized trade on Bitstamp amount to 0.14 USD (354 USD multiplied by the half-spread), an amount which traders may deem negligible. Second, there are frictions beyond the bid-ask spread. For example, trading venues differ in the ways how traders can transfer and withdraw fiat money to and from their accounts. These differences can result in cost and speed differences between the exchanges. Third, not all traders are free to choose where to trade. For example, Bitfinex did not accept US residents as customers during our sample period. Fourth, traders may prefer to trade on a venue in or close to their home country, e.g. because they are more familiar with the legislative regime.

[20] Evidently, ETHSUD is a rather infrequently traded pair on Bitstamp (roughly 5 million transactions over our investigation period, compared to roughly 24 million on Bitfinex and Coinbase Pro) which is why we only have 4410 observations of hourly data on Bitstamp that match the 80% data availability criterion (compared to more than 11,920 one hour intervals that match this criterion on Bitstamp for the pair BTCUSD.

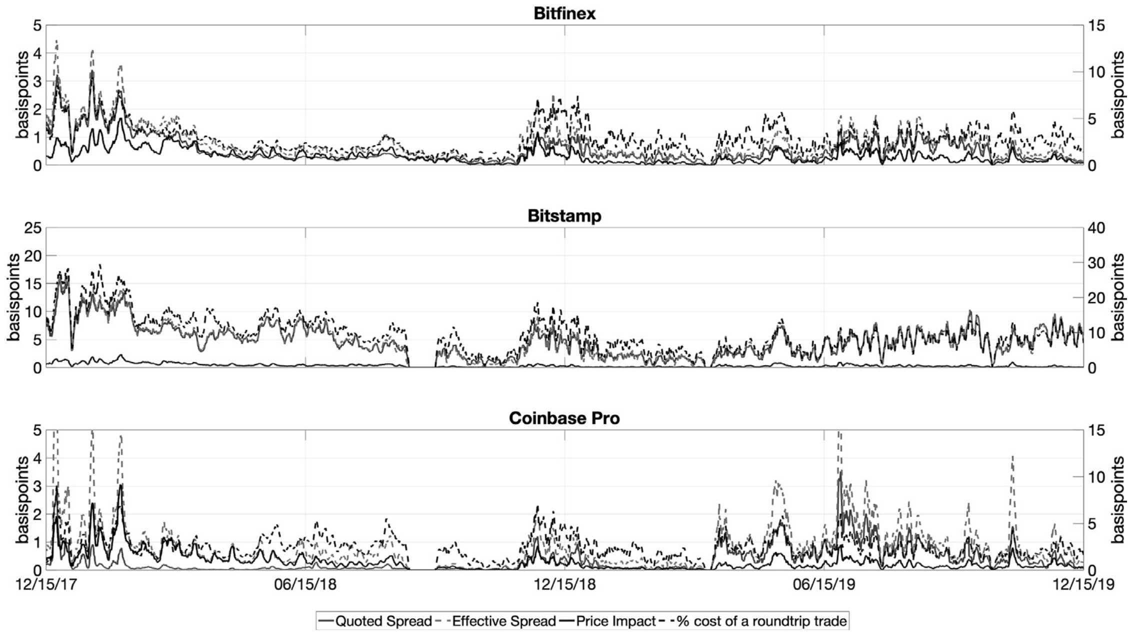

Fig. 1. This figure plots a 72 h moving average of the benchmark liquidity measures for the pair BTCUSD at an hourly resolution. Quoted spread (gray solid line), effective spread (light gray dashed line) and price impact (black solid line) are shown on the left axis, the percentage cost of a roundtrip trade (black dashed line) is shown on the right axis. The unit of measurement is basis points. The sample period is 12/16/2017 to 12/16/2019. Note that for Bitstamp and Coinbase Pro we have a 19 days lack of data in September 2018.

Table 4
Descriptive data for proxy liquidity measures for the pair BTCUSD used in the empirical analysis at a daily resolution. The table reports descriptive statistics for the number of transactions (TX), the dollar volume ($ Vol, million USD), the Amihud measure (Amihud, values *1e6), Roll's returns based measure (Roll_r, basispoints), Roll's price based measure (Roll_p), the Kyle and Obizhaeva measure (Kyle, values *1e3), the Corwin and Schultz measure (CS, basispoints) and the Abdi and Ranaldo measure (AR, basispoints). The sample period is 12/16/2017 to 12/16/2019.

exchange		mean	std. dev.	Q1	median	Q3	no. daily obs
Bitfinex	TX	53,175	45,859	22,622	37,453	67,338	695
	$ Vol	155.1	186.9	38.01	83.25	190.3	
	Amihud	0.015	0.186	0.001	0.002	0.004	
	Roll_r	37.01	45.46	0	25.55	54.57	
	Roll_p	30.71	46.28	0	17.03	40.27	
	Kyle	0.077	0.024	0.060	0.073	0.090	
	CS	21.63	19.49	9.158	15.67	27.45	
	AR	23.60	21.23	10.62	17.02	28.81	
Bitstamp	TX	22,873	18,369	9,705	17,697	30,353	663
	$ Vol	70.44	68.12	24.98	46.50	92.63	
	Amihud	0.038	0.347	0.002	0.003	0.006	
	Roll_r	37.03	45.66	0	25.69	52.68	
	Roll_p	31.14	48.65	0	16.15	40.18	
	Kyle	0.093	0.029	0.073	0.088	0.109	
	CS	24.25	18.12	12.14	19.20	29.73	
	AR	24.06	20.36	10.92	17.89	28.28	
Coinbase Pro	TX	56,928	33,687	35,507	46,277	67,235	668
	$ Vol	91.78	96.97	30.50	61.93	115.1	
	Amihud	0.006	0.059	0.001	0.002	0.003	
	Roll_r	35.58	46.34	0	22.77	51.40	
	Roll_p	29.82	48.65	0	13.73	38.51	
	Kyle	0.085	0.026	0.067	0.081	0.098	
	CS	19.44	16.91	8.527	14.64	24.10	
	AR	22.74	20.02	10.13	16.66	26.74	

price impact is not much larger on Bitstamp than on Bitfinex and Coinbase Pro, implying substantial realized spreads to be earned by liquidity suppliers on Bitstamp.

2.2.2. Transactions-based proxy measures

As noted previously, all transactions-based liquidity measures are calculated from data on open, high, low and closing prices as well as the number of transactions and the dollar trading volume for each subinterval i. The data for the subintervals are then aggregated to one liquidity estimate for each interval t. We use the following transactions-based measures.

- Number of transactions (TX)
  For each interval t we calculate the unweighted average of the number of transactions in the subintervals i, $TX_t = \frac{1}{I}\sum_i TX_{t,i}$, where I denotes the number of subintervals in interval t.
- Dollar Volume ($\$Vol$)

Our second transactions-based measure is the unweighted average of the reported dollar transaction volume $\$Vol_t = \frac{1}{I}\sum_i \$Vol_{t,i}$ in all subintervals i belonging to interval t.

- The Amihud (2002) illiquidity ratio (*Amihud*)
  Amihud (2002) illiquidity ratio for each subinterval is the absolute return (measured from the opening price to the closing price of the subinterval) divided by the dollar trading volume in the subinterval, $Amihud_t = \frac{1}{I}\sum_i \frac{|C_{t,i}/O_{t,i}-1|}{\$Vol_{t,i}}$, where $O_{t,i}$ and $C_{t,i}$ denote the opening and closing price in subinterval i in t, respectively. The illiquidity ratio for interval t is the unweighted average of the ratios for the subintervals i in t. We note that conceptually the illiquidity ratio is a measure of price impact. However, in empirical applications it is routinely used as a proxy for liquidity at large.
- The Roll (1984) serial covariance estimator (*Roll*)
  The Roll (1984) estimator is based on the serial covariance of successive price changes. For each interval t we obtain one spread estimate from the closing prices of all subintervals i in t. If the serial covariance is positive we set the estimator to 0. We calculate two versions of Roll's measure, one based on price changes and one based on returns. The formal expression for the return-based estimator is

$$Roll_t = 2\cdot\sqrt{-\min(cov[\frac{\Delta C_{t,i}}{C_{t,i-1}}, \frac{\Delta C_{t,i-1}}{C_{t,i-2}}], 0)}$$

  where Δ is the first difference operator. In the results section, Roll_p (Roll_r) refers to the price- (return-)based version, respectively.
- The Kyle and Obizhaeva (2016) estimator (*Kyle*)
  Kyle and Obizhaeva (2016) derive an illiquidity index based on the ratio of volatility to dollar volume of an asset within a given interval. It is defined as

$$Kyle_t = \left[\frac{\overline{\sigma^2_{t,i}(r)}}{\sum_i \$Vol_{t,i}}\right]^{1/3}$$

  where the volatility estimator $\overline{\sigma^2_{t,i}(r)}$ is the mean of the squared returns of all subintervals i in interval t.
- The Corwin and Schultz (2012) estimator (*CS*).
  The CS estimator is calculated from the high and low prices of two adjacent subintervals $i, i+1$. It is defined as
  $CS_{i,i+1} = \frac{2(\exp(\alpha)-1)}{1+\exp(\alpha)}$
  $\alpha = \frac{\sqrt{2\beta}-\sqrt{\beta}}{3-2\sqrt{2}} - \sqrt{\frac{\gamma}{3-2\sqrt{2}}}$, $\beta = \left[\ln\left(\frac{H_i}{L_i}\right)\right]^2 + \left[\ln\left(\frac{H_{i+1}}{L_{i+1}}\right)\right]^2$, $\gamma = \left[\ln\left(\frac{H_{i,i+1}}{L_{i,i+1}}\right)\right]^2$
  H_i and L_i denote the high and low prices, respectively, in subinterval i, while $H_{i,i+1}$ and $L_{i,i+1}$ refer to the high and low price, respectively, of two adjacent subintervals i and $i+1$. We follow Corwin and Schultz (2012) and set negative values of the proxy to zero. The CS_t estimator for period t is the unweighted average of all CS estimators for adjacent subintervals in t.
  Corwin and Schultz (2012) propose a method to adjust their estimator for the overnight trading halt. We do not need to implement this modification because cryptocurrency exchanges operate 24 h a day and seven days a week. There are thus no regular trading halts.
- The Abdi and Ranaldo (2017) estimator (*AR*)
  Abdi and Ranaldo (2017) propose an estimator based on the natural logarithms of high, low and closing prices in subinterval i, denoted $h_i = \ln(H_i)$, $l_i = \ln(L_i)$ and $c_i = \ln(C_i)$, respectively. Further, denote by $\bar{p}_i = (h_i + l_i)/2$ the midpoint between the high and the low log prices in subinterval i. We use the 'two-day corrected' version of the estimator which uses high and low price data from two adjacent subintervals i and $i+1$. It is defined as
  $AR_i = \sqrt{\max\{4(c_i - \bar{p}_i)(c_i - \bar{p}_{i+1}), 0\}}$
  The AR_t estimator for interval t is the average of the $AR_{t,i}$ measures for all adjacent subintervals i in t,
  $AR_t = \frac{1}{I-1}\sum_{i=1}^{I-1} AR_{t,i}$.

Table 4 reports descriptive statistics for our proxy liquidity measures at the daily frequency (results at the hourly and 15-day frequency are again available upon request). When interpreting the numbers it should be kept in mind that several measures (Abdi and Ranaldo, 2017; Corwin and Schultz, 2012; Kyle and Obizhaeva, 2016; Roll, 1984) estimate the effective bid-ask spread while the Amihud (2002) illiquidity ratio is a measure of price impact and the two volume metrics measure the number of trades and the dollar trading volume, respectively. However, even among those measures that estimate the effective spread there are large differences. The Roll (1984) estimator delivers the largest and the Kyle and Obizhaeva (2016) estimator delivers the smallest spread estimates. We will compare the mean values shown in Table 4 to the effective spread calculated from high-frequency quote data in Section 3.5 below.

For the currency pair ETHUSD descriptive statistics for our proxy liquidity measures are reported in Table 10 in the appendix. As for our benchmark measures the results indicate that the pair ETHUSD is less liquid than BTCUSD: volume-based proxy measures show lower values, while price-based measures are higher.[21] We will discuss these results in more detail in Section 3.7.

3. Results

We present the results in six steps. We first report time-series correlations between the transactions-based proxies and the benchmark measures. Correlations are a global measure of linear dependence. To analyze whether the dependence structure is different in the tails of the distribution we analyze, in step 2, quantile dependencies based on the empirical distribution functions. In a third step we aggregate the transactions-based measures into a composite measure, the first principal component of the transactions-based measures, and analyze whether it has higher time-series correlation with the benchmark measures than the best of the individual transactions-based measures. In order to investigate whether the performance of the transactions-based measures depends on the specific market environment we then, in step 4, split our sample along several dimensions. Specifically, we estimate time-series correlations between the proxies and the benchmark measures for the first and the second half of our sample period, for high and low return periods, high and low volatility periods and high and low volume periods. Next we report the mean absolute errors and root mean squared errors of the transactions-based measures. Finally, to capture the cross-sectional dimension, we analyze how frequently the liquidity ranking across the exchanges produced by the transactions-based measures is equal to the ranking produced by the benchmark measures.

In Sections 3.1–3.6 we present results for the currency pair BTCUSD in detail; qualitative results for the pair ETHUSD are similar in most respects and are summarized in Section 3.7.

3.1. Time-series correlations

An accurate transactions-based measure should capture the time-series variation in liquidity and should thus be positively cor-

[121] We note that the price-based Roll measure delivers an estimate of the dollar spread, not of the percentage spread. The numerical values are lower for ETHUSD than for BTCUSD because the dollar price of ethereum is only a fraction of the bitcoin price.

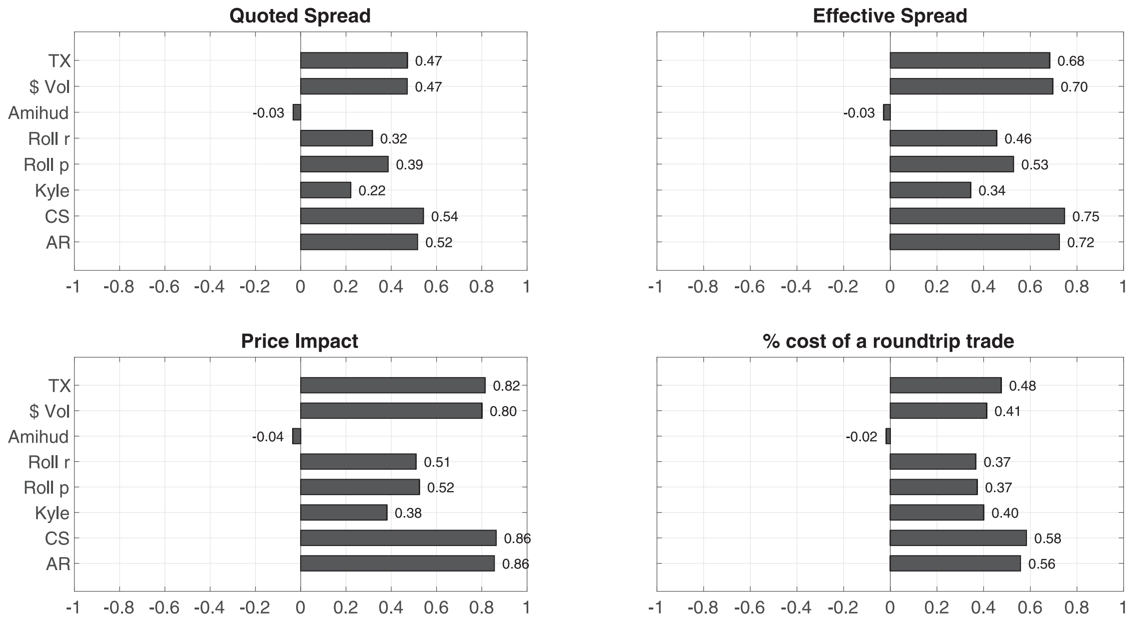

Fig. 2. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

related with the benchmark measures. We therefore estimate time-series correlations between the low-frequency measures and the high-frequency measures. We do so separately for three exchanges (Bitfinex, Bitstamp and Coinbase Pro) and three time frames (1-minute data (1-hour-data, daily data) aggregated to the hourly (daily, 15-daily) frequency). As it turns out, the results for the three exchanges are very similar. We therefore report averages across the trading venues.[22] In the description of the results we emphasize the findings for the daily data (i.e. hourly data aggregated to the daily frequency). We believe that most researchers using transactions-based liquidity proxies will do so to obtain daily estimates, and the hourly raw data required to calculate these daily estimates are easily and freely available, e.g. from cryptodatadownload.com.

The results are presented in Figs. 2–4. Focusing on the results for the daily intervals (Fig. 2) we find that the Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators perform best. The Corwin and Schultz (2012) estimator exhibits the highest time series correlation and the Abdi and Ranaldo (2017) estimator the second-highest correlation with the quoted spread, the effective spread and the cost of a roundtrip trade, with correlations ranging from 0.54 to 0.75. The correlation with the price impact is markedly higher, at 0.86 for both the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimator.

Interestingly, the number of transactions and the dollar trading volume are highly correlated with the four benchmark measures, and particularly so with the price impact (with correlation coefficients of 0.82 and 0.80 for the number of transactions and the dollar volume, respectively. What is most surprising is the sign of the coefficients. Both measures are *positively* related to the benchmark measures, implying that higher trading activity is associated with higher execution costs.

The other transactions-based proxies achieve much lower correlations. The measure that performs worst in our horse race is the Amihud (2002) illiquidity ratio. It is virtually uncorrelated with the benchmark measures, and the sign of the correlation is even negative.[23] The poor performance of the Amihud (2002) illiquidity ratio deserves discussion because this ratio is widely used as a measure of liquidity in empirical microstructure research. We argue that the lack of correlation between the illiquidity ratio and the benchmark measures is caused by the strong and positive relation between liquidity and trading activity discussed above. The illiquidity ratio is based on the presumption that, in a less liquid market, a given dollar trading volume will have a larger impact on prices and will thus result in a larger price change. Put differently, for a given price change higher volume points to a more liquid market and should thus be associated with lower execution costs according to the inherent logic of the measure. However, in the markets under investigation volume is *positively* related to execution costs, a relation that runs counter the logic of the illiquidity ratio. We wish to reemphasize that the finding of a positive relation between trading activity and execution costs, even though at odds with the predictions of standard theory, is not confined to the cryptocurrency markets under consideration here. Bogousslavsky and Collin-Dufresne (2020) have recently documented a similar finding for large US stocks.

The results for the two alternative time frames (one-minute data aggregated to the hourly frequency (Fig. 3) and daily data aggregated to the 15-day frequency (Fig. 4) are similar. The Corwin and Schultz (2012) and the Abdi and Ranaldo (2017) estimators yield higher correlations with the benchmark measures than the other transactions-based estimators. The strong and positive correlations documented above for the daily data frequency between the transaction frequency and dollar trading volume on the one hand and the benchmark measures on the other hand persist at the other data frequencies. In fact, at the 15-daily frequency the volume-based proxies perform better than the other proxies for three out of four benchmark measures.

The performance of the Kyle and Obizhaeva (2016) estimator is better at higher data frequencies while the Roll (1984) estimator appears to perform better at lower frequencies. The

[22] Results for individual exchanges are available upon request.

[23] In their study on the foreign exchange market Karnaukh et al. (2015) find that the Amihud (2002) illiquidity ratio performs reasonably well. When calculating the illiquidity ratio the authors use the number of transactions as a proxy for the dollar trading volume because they do not have access to volume data.

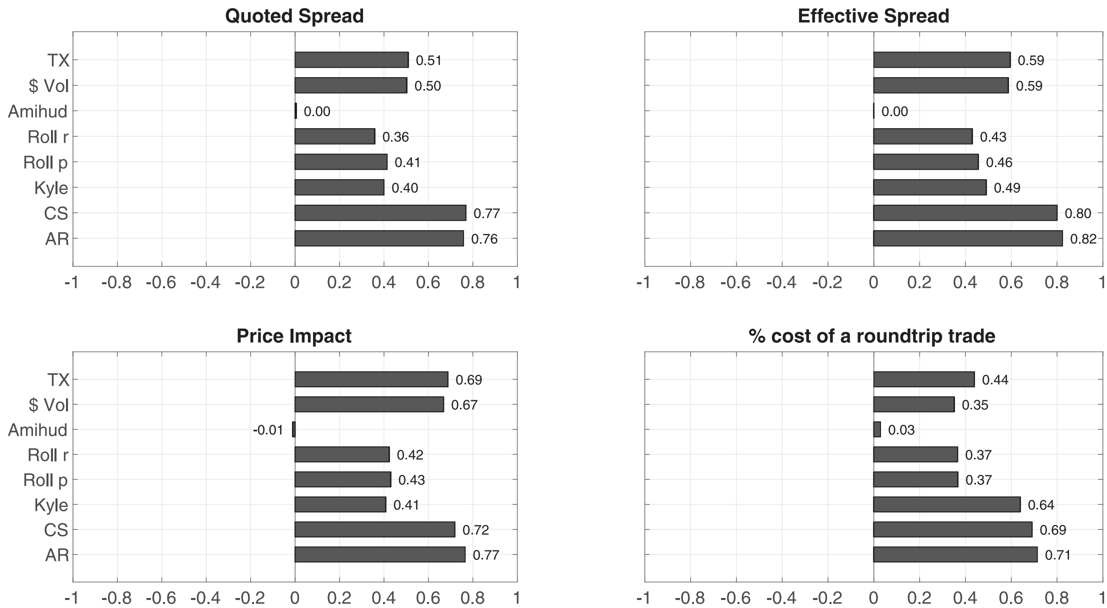

Fig. 3. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on an hourly basis over the sample period 12/16/2017 to 12/16/2019.

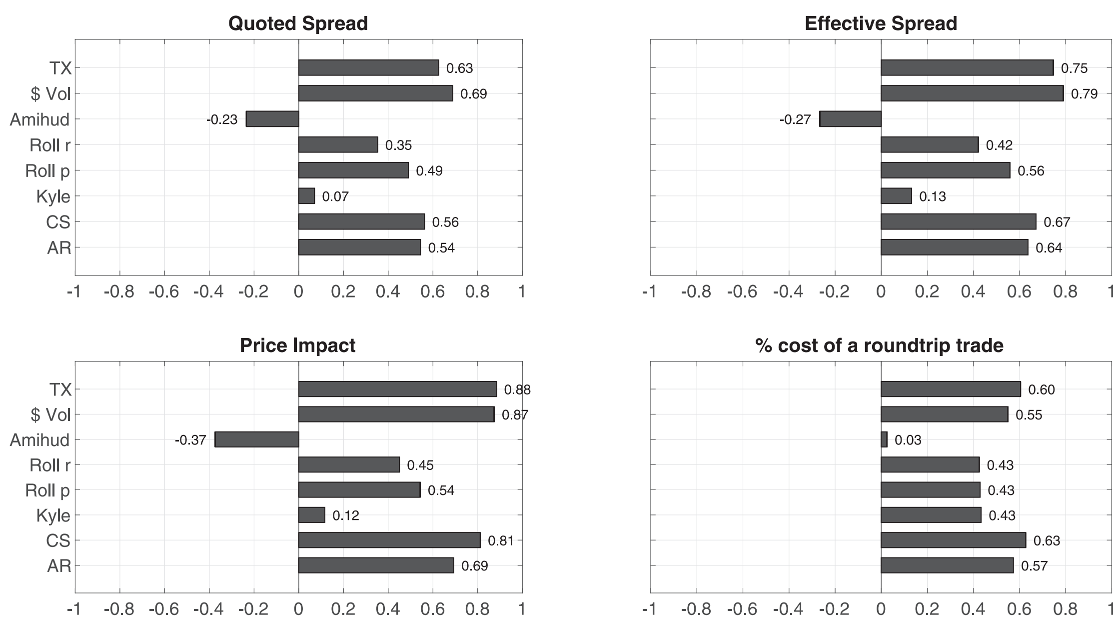

Fig. 4. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a 15-daily basis over the sample period 12/16/2017 to 12/16/2019.

Amihud (2002) illiquidity ratio continues to be the worst-performing measure. As explained above, the most likely reason is the positive relation between trading activity and spreads in the cryptocurrency markets.

So far we have documented differences in correlations across the different transactions-based measures, but we do not know whether the differences are significant. We therefore now perform a formal test based on the Fisher r-to-z transformation. Specifically we test, separately for each of the four benchmark measures and the three trading venues, whether the correlation between the best-performing proxy and the benchmark measure is significantly higher than the correlation between the second-best performing proxy and the benchmark. The results are reported in Table 5. The evidence in favor of significant differences is limited to the two highest data frequencies. The Corwin and Schultz (2012)) and the Abdi, Ranaldo (2017) estimators perform best. Each of them significantly outperforms the second-ranking estimator in six cases. At the 15-daily data frequency there is no compelling evidence in favor of significant differences between the two best-performing transactions-based measures.

3.2. Quantile dependence

The correlation between the time series of transactions-based proxies and the benchmark measures of liquidity provides a global measure of dependence. However, it is conceivable that a proxy measure that fits the benchmark well in times of high liquidity (i.e. in times of low bid-ask spreads) performs poorly in times of low liquidity and vice versa. We therefore use quantile dependence

Table 5
This table reports the best performing (in terms of correlation) proxy liquidity measure for each of the four benchmark liquidity measures (second best in parentheses) for the pair BTCUSD. **/* denote statistical significance of the difference between the best and second best measure's correlation coefficient using Fisher r-to-z transformation at the 1%/5% level. For example, the entry CS (AR)** for hourly data of Bitfinex in column QS indicates that CS features the highest correlation with the quoted spread, while AR is ranked second. These correlations are statistically different at the 1% level. The sample period is 12/16/2017 to 12/16/2019.

		QS	ES	PI	CRT
hourly	Bitfinex	CS (AR)**	AR (CS)**	AR (CS)**	Kyle (AR)
	Bitstamp	AR (CS)**	AR (CS)**	TX ($Vol)**	AR (CS)**
	Coinbase Pro	CS (AR)**	CS (AR)**	AR (CS)**	CS (AR)
daily	Bitfinex	CS (AR)	CS (AR)	CS (TX)	CS (AR)
	Bitstamp	CS (AR)*	CS (AR)*	CS (AR)	CS (AR)*
	Coinbase Pro	$Vol (CS)	$Vol (TX)**	AR (CS)	CS (AR)
15-daily	Bitfinex	TX ($Vol)	TX ($Vol)	TX ($Vol)	CS (AR)
	Bitstamp	$Vol (TX)	$Vol (TX)	$Vol (TX)	TX ($Vol)
	Coinbase Pro	Roll_p ($Vol)	$Vol (Roll_p)	$Vol (TX)	CS (AR)

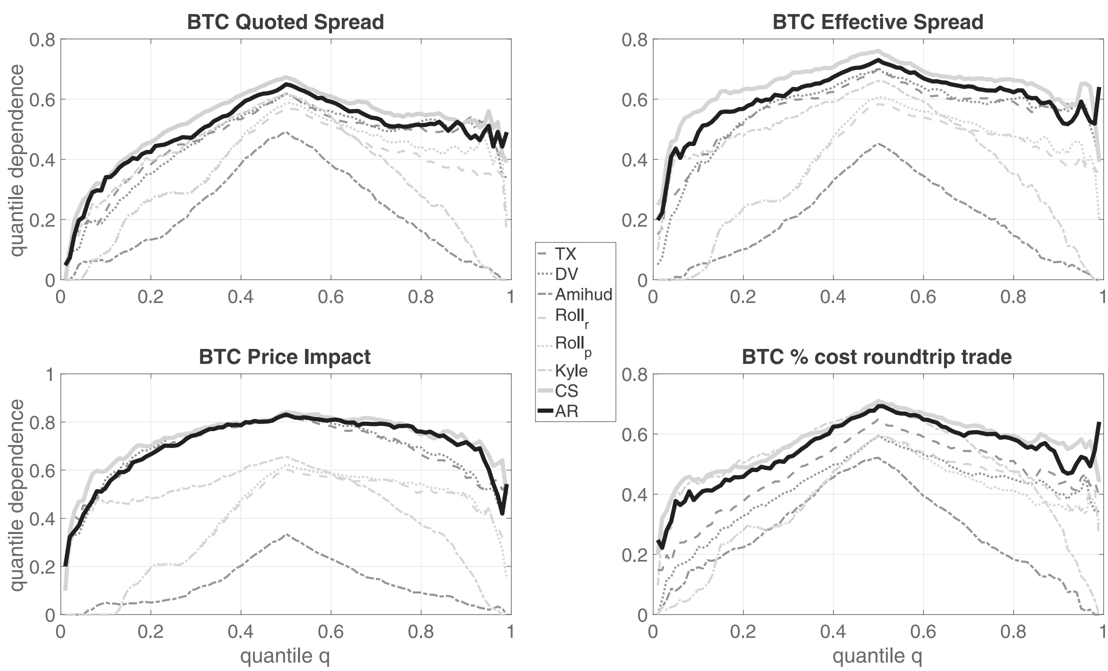

Fig. 5. This figure plots the quantile dependence (averaged across the three trading venues) as a function of the quantile q in steps of 0.01 for the four benchmark liquidity measures and each of the 8 proxy measures for the pair BTCUSD for daily data. The sample period is 12/16/2017 to 12/16/2019.

to analyze the dependence structure between the benchmark and proxy measures in more detail. The quantile dependence of order q between two random variables η_p and η_b is generally defined as the conditional probability that $F_p(\eta_p)$ is smaller (greater) than q given that $F_b(\eta_b)$ is smaller (greater) than q for $q \le 0.5$ ($q > 0.5$):[24]

$$\lambda_q^{p,b} = \begin{cases} P[F_p(\eta_p) \le q \mid F_b(\eta_b) \le q], & \text{for } q \in (0, 0.5] \\ P[F_p(\eta_p) > q \mid F_b(\eta_b) > q], & \text{for } q \in (0.5, 1). \end{cases}$$

In our application the subscripts b and p refer to the benchmark measures and the transactions-based proxies, respectively. q denotes a quantile and F denotes the cumulative distribution function (CDF). An empirical estimate of $\lambda_q^{p,b}$ is given by

$$\widehat{\lambda}_q^{p,b} = \begin{cases} \dfrac{1}{Tq} \sum_{t=1}^{T} \mathbf{1}_{[\widehat{F}_p(\widehat{\eta}_{p,t}) \le q, \widehat{F}_b(\widehat{\eta}_{b,t}) \le q]}, & \text{for } q \in (0, 0.5], \\ \dfrac{1}{T(1-q)} \sum_{t=1}^{T} \mathbf{1}_{[\widehat{F}_p(\widehat{\eta}_{p,t}) \le q, \widehat{F}_b(\widehat{\eta}_{b,t}) \le q]}, & \text{for } q \in (0.5, 1). \end{cases}$$

$\hat{F}_j\left(\hat{\eta}_j\right)$; $j \in \{b, p\}$ denotes the empirical distribution functions of the benchmark and proxy measure, respectively. We estimate it using scaled ranks, i.e. we transform the data into ranks and then rescale these ranks onto the unit interval.

Intuitively, quantile dependence works as follows. For any $q \le 0.5$ consider the $q \cdot T$ smallest observations for the benchmark measure, where T is the total number of observations. Then consider the $q \cdot T$ smallest values for a transactions-based proxy and determine the fraction of coinciding values. This fraction is the estimate of the quantile dependence, $\widehat{\lambda}_q^{p,b}$.

We use data at the daily frequency[25] to estimate the quantile dependence separately for each trading venue and then calculate averages across venues. We present the results using quantile dependence plots which show the quantile dependence as a function of q. Higher quantile dependence implies a closer relation between the benchmark measures and the transactions-based proxies. The results for our four benchmark measures and eight proxies are shown in Fig. 5. The dependence between the benchmark and proxy measures is generally stronger in the center of the distribution and weaker in the tails. The dependence in the tails ap-

[24] See e.g. Duan et al. (2019).

[25] Results for the hourly frequency are qualitatively similar and are available upon request. The number of observations at the 15-day frequency is too low to reliably estimate quantile dependence, particularly in the tails of the distributions.

Table 6
This table reports mean correlations among the first principal component of proxy liquidity measures and the four benchmark measures for the pair BTCUSD. Column 'expl. var' shows the percentage of total variance explained by the first principal component. Rows denoted 'mean' contain equally-weighted averages across exchanges. The sample period is 12/16/2017 to 12/16/2019.

		expl. var	QS	ES	PI	CRT
hourly	Bitfinex	56.34	0.761	0.707	0.773	0.576
	Bitstamp	49.75	0.725	0.789	0.718	0.724
	Coinbase Pro	57.15	0.582	0.815	0.780	0.630
	mean	54.41	0.689	0.770	0.757	0.643
daily	Bitfinex	56.38	0.818	0.741	0.859	0.657
	Bitstamp	57.87	0.702	0.760	0.885	0.767
	Coinbase Pro	57.87	0.078	0.764	0.844	0.253
	mean	57.37	0.533	0.755	0.863	0.559
15-daily	Bitfinex	53.10	0.765	0.787	0.896	0.539
	Bitstamp	54.09	0.702	0.724	0.875	0.723
	Coinbase Pro	56.78	0.489	0.778	0.810	0.594
	mean	54.65	0.652	0.763	0.860	0.619

pears to be asymmetric, it is higher for larger than for smaller values.[26]

With respect to the ranking of the transactions-based liquidity measures the results from the quantile dependence analysis are consistent with the results shown in Fig. 2 above. In particular, the Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators perform very well over the entire distribution, i.e. for high as well as low levels of liquidity. The two measures of trading activity, the number of transactions and the dollar volume, also perform well. As before, the Amihud (2002) illiquidity ratio performs poorly.

3.3. Composite estimator

The different transactions-based measures capture different aspects of liquidity. It is, therefore, conceivable that a combination of these measures better captures the time-series variation of liquidity. To test whether this is the case we construct a composite estimator based on the eight low-frequency measures. We first standardize all variables by subtracting the mean and dividing by their standard deviation. We then extract the first principal component of the standardized data and estimate the time-series correlations between the first principal component and the benchmark measures. Table 6 shows the results for each time frame and each of the three exchanges.

The first principal component explains roughly 55% of the variation in the data. The time-series correlations are highest with the effective spread and the price impact, with average values (across the three exchanges) ranging from 0.76 to 0.86. The time-series correlations are lower (with values ranging from 0.53 to 0.69) when one of the other benchmark measures is used.

Comparing the results in Table 6 to those in Figs. 2–4 reveals that the best performing individual estimators achieve higher time-series correlations than the composite estimator for each benchmark measure at the hourly and 15-daily frequencies. At the daily frequency the composite estimator performs virtually equally well as the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators. We thus conclude that the benefit of calculating all transactions-based proxies and aggregating them to a composite estimator is limited, and is confined to specific data frequencies.

[26] When interpreting the results note that a quantile dependence of 0.5 for $q = 0.5$ is expected when the two distributions are independent.

3.4. Sample splits

It may be the case that some of the transactions-based liquidity measures perform better under specific circumstances, e.g. earlier or later in the sample period or at times of high or low volatility. To shed light on this issue we split our sample along several dimensions. The analysis is performed using the hourly data aggregated to the daily level.[27] We start by separately considering the first and the second half of the sample period, with the resulting sub-samples covering 12/16/2017 00:00 UTC to 12/15/2018 24:00 UTC and 12/16/2018 00:00 UTC to 12/16/2019 00:00 UTC, respectively. Subsequently we split the sample into terciles according to the signed return (measured by the average of the one-hour returns within a daily interval), return volatility (measured by the standard deviation of the hour-by-hour returns within a one-day interval) and the dollar volume. We then calculate separate time-series correlations for the first and the third tercile.

For each sample split we present results for all four benchmark measures. We first calculate time-series correlations for each trading venue and then average the correlations across venues. These average correlations are shown in Figs. 6–9. During the second half of the sample period the correlations between the transactions-based proxies and the benchmark measures are lower than those in the first half for most proxies (Fig. 6). This reduction is much more pronounced for the quoted spread and the cost of a roundtrip trade than for the effective spread and the price impact. These results allow the conclusion that the performance of the transactions-based proxies is sufficiently stable over time when they are used to track the time-series variation of the effective spread.

The time series correlations between the transactions-based proxies and the benchmark measures do not differ much between high and low return periods (Fig. 7). The performance tends to be better in low return periods.

When we consider the sub-samples split by volatility and volume (Figs. 8 and 9, respectively) we find that the time series correlations are higher in high volatility and high volume periods for all benchmark measures.

What is most important, though, is that our previous results concerning the relative performance of the low-frequency liquidity proxies still hold. The Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators have the best overall performance under almost all conditions. The Kyle and Obizhaeva (2016) estimator performs rather well in the low volume periods but is unable to track liquidity across high volatility periods.

3.5. Mean absolute errors and root mean squared errors

The previous analyses have focused on the ability of the transactions-based proxies to capture the time-series variability of the benchmark liquidity measures. An alternative question is whether the proxy measures are able to accurately estimate the level of the benchmark measures. Investors take the level of liquidity into account in their trading strategies and portfolio allocations. Further, because the magnitude of the execution costs determines whether a given price difference (e.g. for the same cryptocurrency at two different trading venues) can be profitably exploited, the level of liquidity is also related to market efficiency.

We use as performance metrics the prediction error between the liquidity benchmark and the liquidity proxy as measured

[27] Results for the hourly frequency are qualitatively similar and are available upon request. The number of observations at the 15-day frequency is too low to split the sample into terciles.

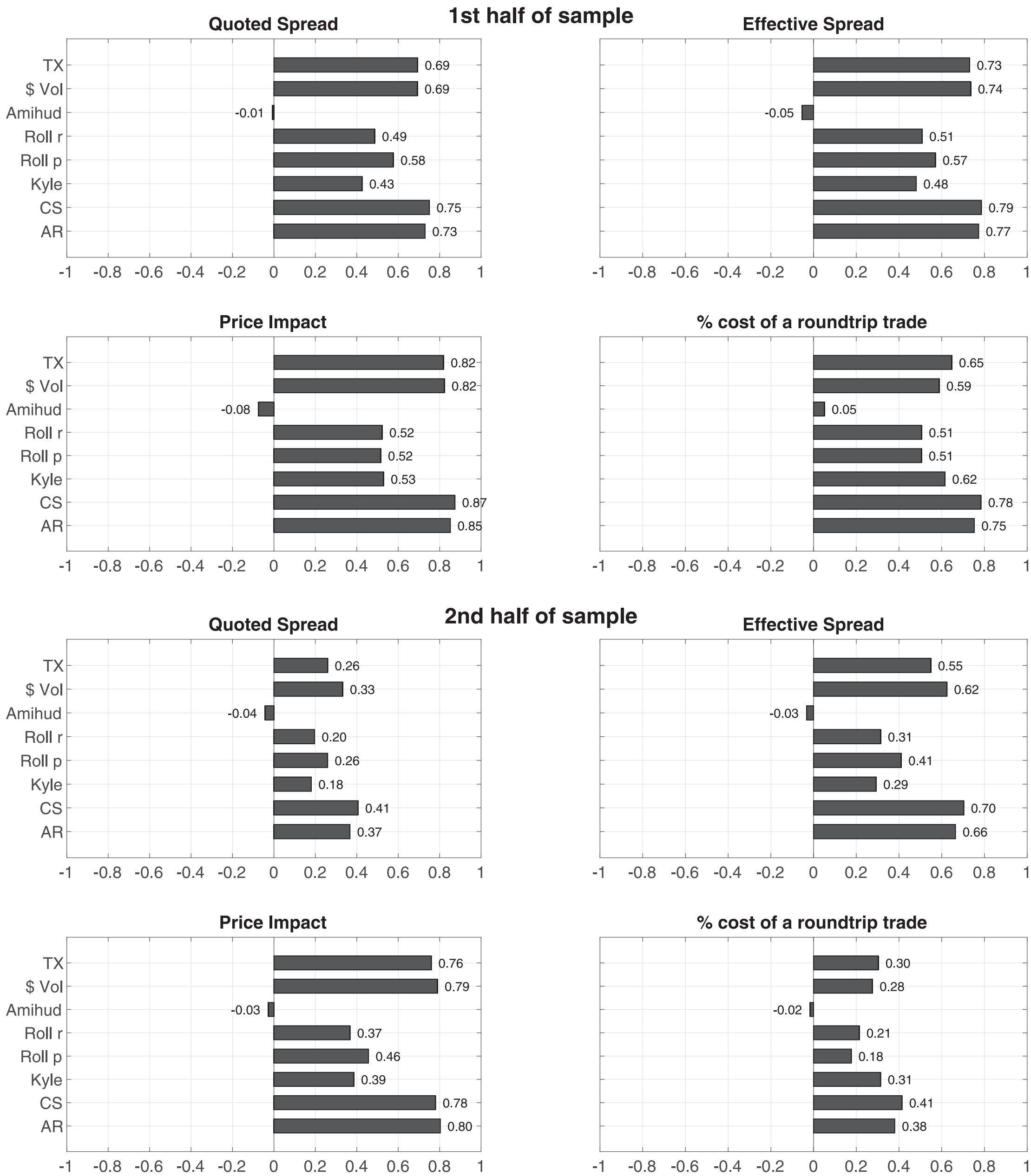

Fig. 6. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the first and second half of the sample for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

by the root mean squared error (RMSE) and the mean absolute error (MAE). Table 7 reports the corresponding results for five transactions-based measures (Amihud, Roll_r, Kyle, CS and AR) with respect to the two benchmark measures *ES* and *PI*. We exclude as proxies the trading frequency, Roll's price based estimator and the dollar volume from the analysis because these measures are obviously unable to directly capture the percentage transaction costs.

When it comes to capturing the level of the effective spread the Kyle and Obizhaeva (2016) estimator performs very well, particularly when data at lower frequencies are used. It has the lowest RMSE and MAE for two data frequencies (daily, and 15-daily). Surprisingly, the Amihud (2002) illiquidity ratio comes close to Kyle and Obizhaeva (2016). The Corwin and Schultz (2012) and Abdi and Ranaldo (2017) proxies capture the level of the percentage effective spread very well at the highest data frequency but very poorly at lower frequencies. The Roll (1984) measure performs worst.

When the price impact is used as benchmark measure the Amihud (2002) illiquidity ratio and the Kyle and Obizhaeva (2016) estimator yield the best results for all frequencies and both metrics for the prediction errors. The Roll (1984) estimator and the two measures based on high and low prices, the Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators, are unable to capture the levels of the benchmark measures.

One striking observation is that the levels of the Roll (1984), Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators appear to strongly depend on the data frequency. All three measures deliver values which increase strongly as we move to lower

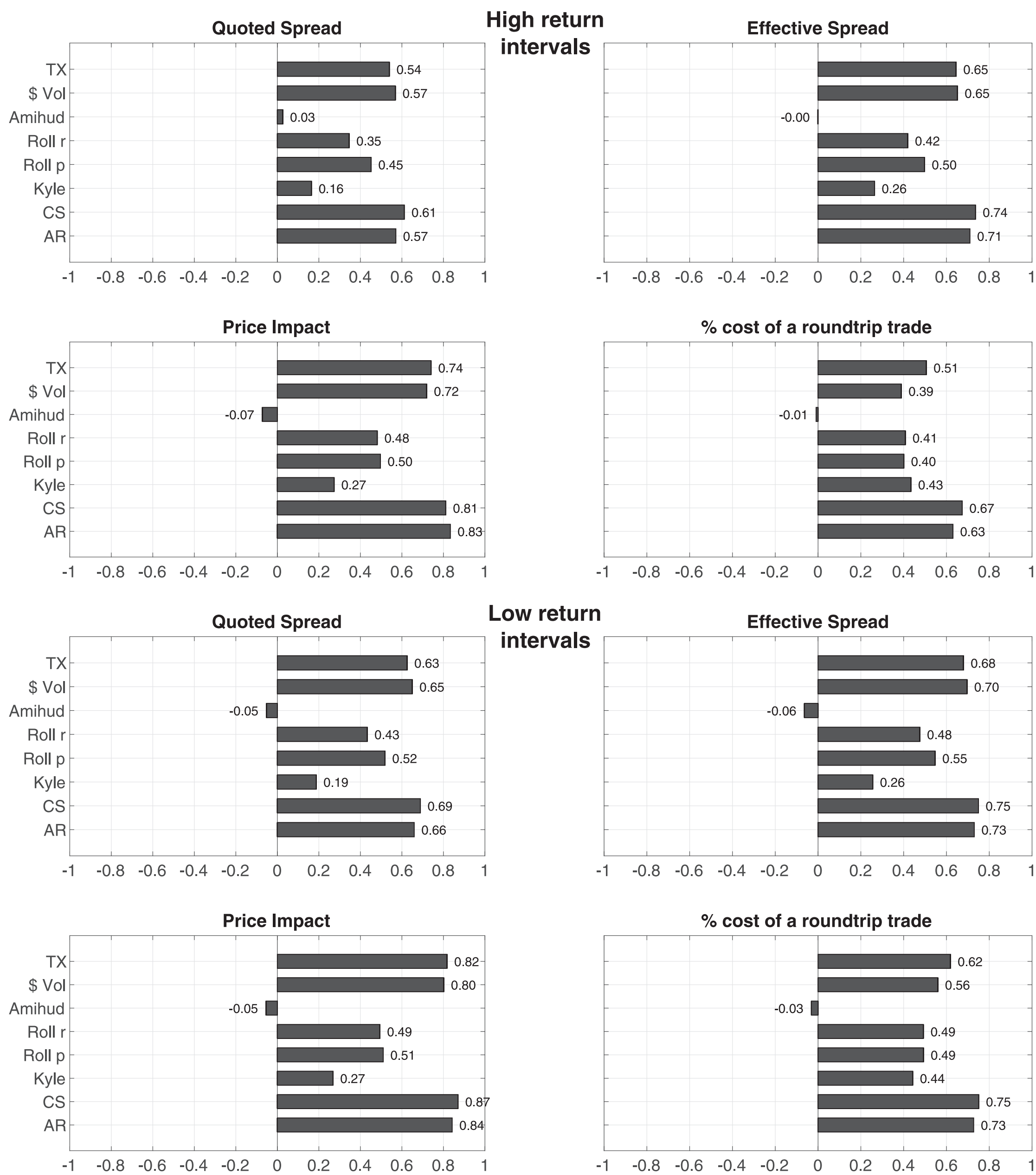

Fig. 7. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low return intervals for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

data frequencies. This phenomenon as such is not new. Already in his original paper, Roll (1984) obtained much larger spread estimates from weekly than from daily data. Harris (1990) argued that the difference can partly be explained by a small sample bias in the estimator of the serial covariance.[28] The results in Table 7 suggest that the high-low spread estimators developed by Corwin and Schultz (2012) and Abdi and Ranaldo (2017) are subject to a similar bias.

[28] Specifically, he showed that the expected value of the serial covariance estimator is $E(SCov) = \frac{-s^2}{4} - \frac{-\sigma^2}{n}$ where s is the spread, σ^2 is the variance of price changes and n is the number of observations. The bias in the serial covariance estimator, $-\frac{-\sigma^2}{n}$, increases with the square of the observation interval. Under ideal conditions (i.e. i.i.d. returns and continuous trading seven days a week, as is the rule in cryptocurrency markets), the variance of weekly price changes is seven times the variance of daily price changes while the number of observations is one seventh. Consequently, the bias in weekly data is 49 times the bias in daily data.

3.6. Cross-sectional analysis

One potential application of transactions-based liquidity measures is to compare the liquidity of different trading venues. A good proxy measure should produce the same ranking of the venues as the benchmark measures. Therefore, in order to evaluate the low-frequency measures we simply analyze how frequently the liquidity ranking across trading venues produced by the transactions-based measures is equal to the ranking produced by the benchmark measures. We perform the analysis separately for each exchange pair (Bitfinex/Bitstamp, Bitfinex/Coinbase Pro, and Bitstamp/Coinbase Pro) and for each time frame. For each interval (one hour, one day, 15 days) and each exchange pair we record the corresponding liquidity ranking based on the benchmark measures and based on the transactions-based proxies and then simply count the fraction of identical rankings. By chance, this

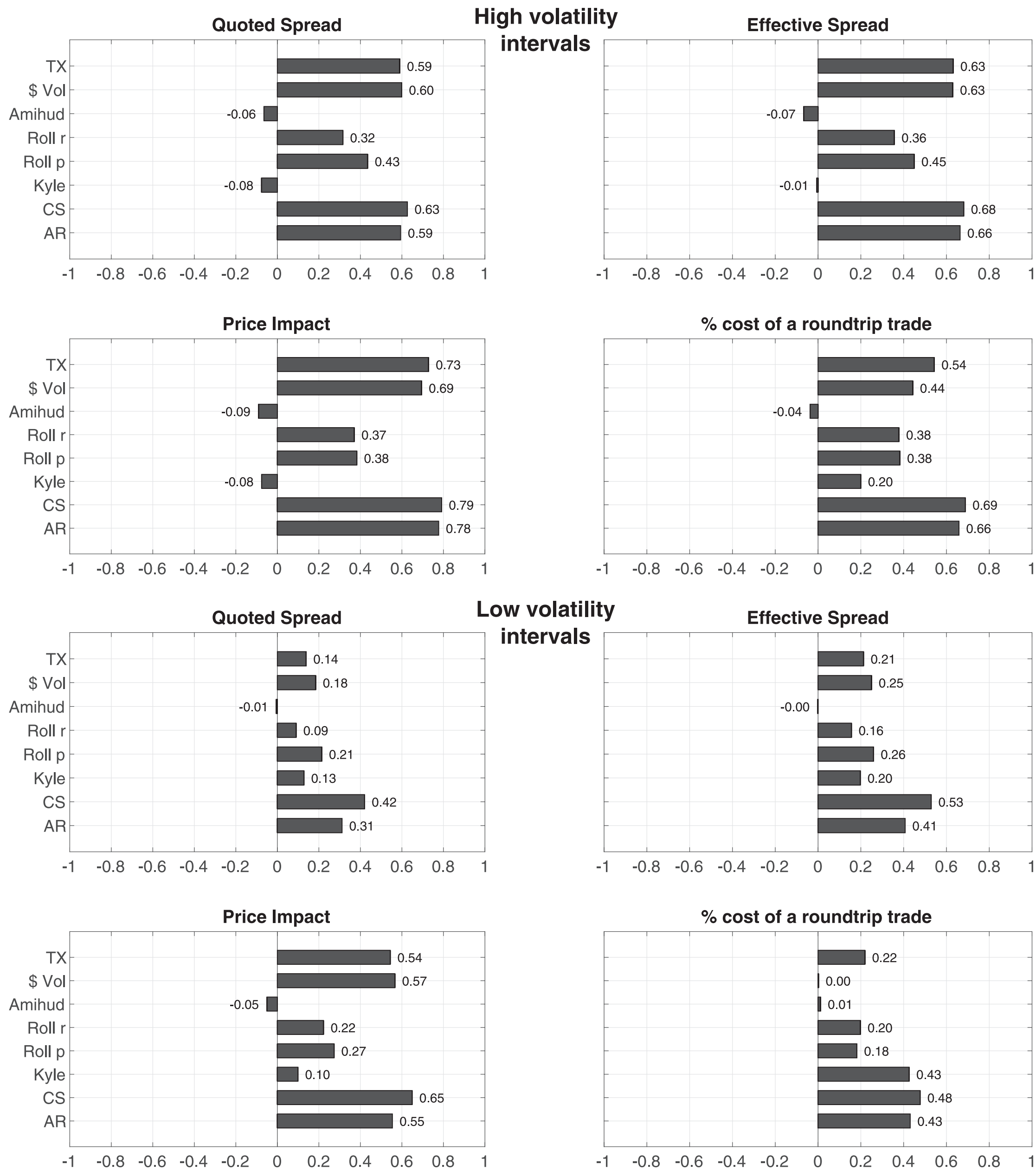

Fig. 8. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low volatility intervals for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

fraction should be 50%. Therefore, we test whether the actual fractions are significantly larger than 50% using a simple binomial test. The results are presented in Table 8.

Two general patterns emerge. First, the results for the quoted spread, the effective spread and the cost of a roundtrip trade are better than those for the price impact. Several of the transactions-based measures (in particular the Amihud (2002) illiquidity ratio, the Kyle and Obizhaeva (2016) measure, and the Abdi and Ranaldo (2017) and Corwin and Schultz (2012) estimators) replicate the ranking of trading venues according to these three benchmark measures well, with fractions of correct rankings ranging up to 97.3%. For the price impact, on the other hand, the percentage of matching rankings is lower and is often close to 50% even for the best-performing estimators. The liquidity proxies thus do not provide valuable information on the ranking of price impacts across different trading venues.

Second, the ability of some transactions-based measures to capture the cross-venue differences in liquidity depends on the data frequency. In particular, the higher the data frequency (and, correspondingly, the number of observations), the better the performance of the Abdi and Ranaldo (2017) estimator. The Amihud (2002) illiquidity ratio, the Corwin and Schultz (2012) and the Kyle and Obizhaeva (2016) estimators, in contrast, are more consistent. They perform well at all data frequencies.[29]

Overall, when considering all 36 comparisons (4 benchmark measures, 3 trading venue pairs, 3 data frequencies) of each

[29] Note that the Amihud (2002) illiquidity ratio does not capture the liquidity differences between Bitfinex and Coinbase well. However, as documented in Table 3 above, the liquidity differences between these two exchanges are small. When liquidity differences are small, ranking venues according to their liquidity is less important.

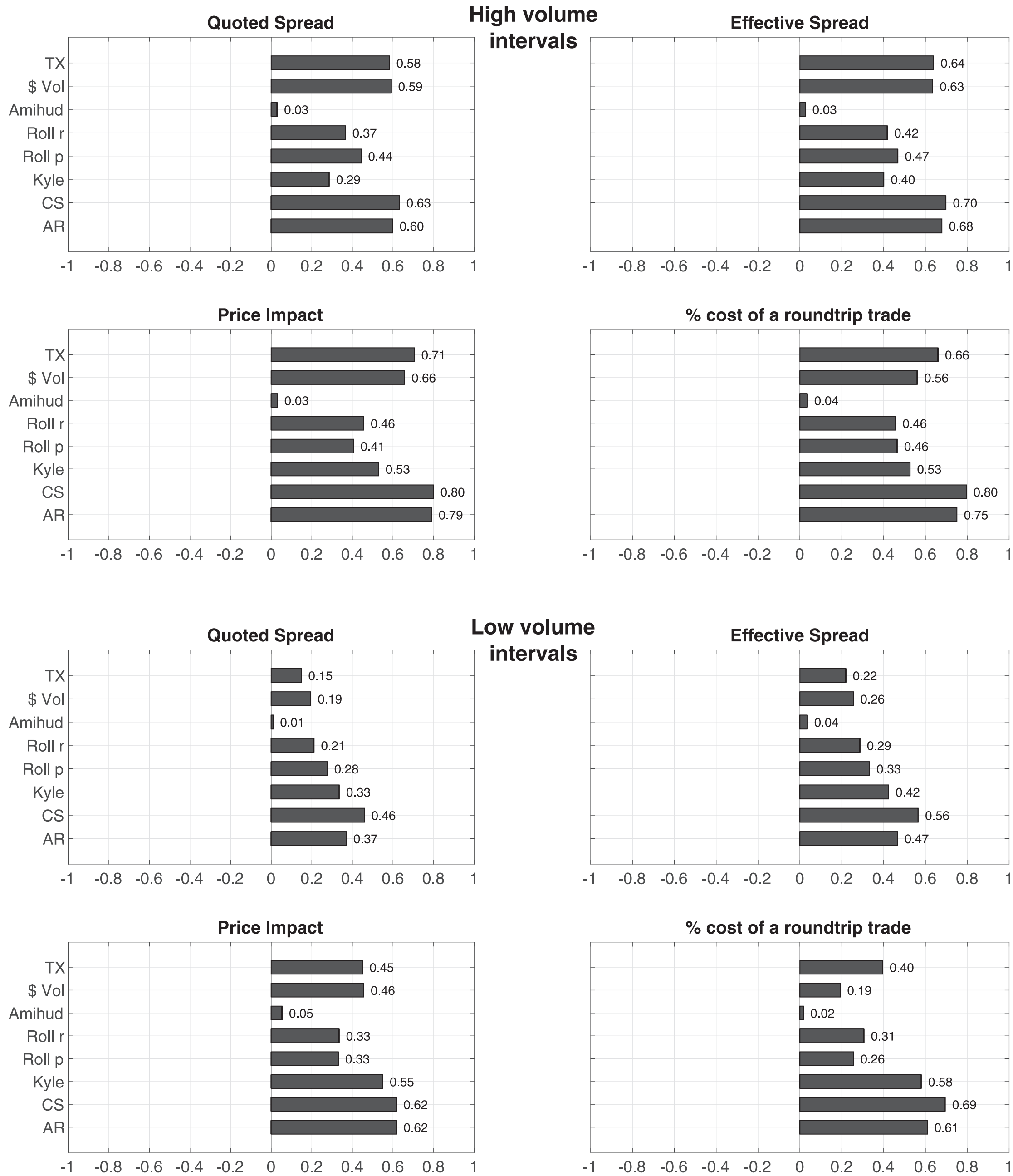

Fig. 9. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low volume intervals for the pair BTCUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

transactions-based measure to the benchmark measures, the Corwin and Schultz (2012) estimator achieves a rate of correct rankings significantly more frequently than expected by pure chance (i.e., a rate of correct rankings significantly above 50%) in 26 cases. The corresponding numbers for the Amihud (2002) illiquidity ratio, the Kyle and Obizhaeva (2016) measure and the Abdi and Ranaldo (2017) estimator are 26, 25 and 18, respectively.

3.7. Results for ethereum

In this section we briefly discuss our results for the currency pair ETHUSD which are in most respects qualitatively similar to those for BTCUSD. As for BTCUSD, results for the three trading venues are very similar. We therefore report averages across the venues. All tables and figures we are referring to are in the appendix.

When considering the time-series correlations between our low-frequency liquidity measures and the high-frequency benchmark measures we mostly find correlation levels that are slightly lower for ETHUSD than for BTCUSD, particularly at the lowest data frequency (see Figs. 10–12). The Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators yield the highest correlations for *QS*, *ES* and *PI* at the hourly and the daily frequency, with often almost identical correlation levels achieved by these two estimators. When the cost of a roundtrip trade $CRT(Y)$[30] is used

[30] We set the dollar trading volume Y to USD 17,400 which corresponds to the 99% quantile of the aggregate trade size distribution for the currency pair ETHUSD.

Table 7
This table reports average root mean squared errors (RMSE) and average mean absolute errors (MAE) for the three exchanges for the pair BTCUSD. The panels in the table refer to hourly, daily and 15-daily results respectively. Lines depict proxy measures whilst columns in the table refer to the benchmark liquidity measures effective spread (ES) and price impact (PI). Note that Amihud is conceptually a proxy for the price impact, while the other low-frequency measures Roll_r, Kyle, CS and AR estimate the spread. All values are multiplied by 1000. The sample period is 12/16/2017 to 12/16/2019.

	RMSE		MAE	
Hourly	ES	PI	ES	PI
Amihud	0.487	0.292	0.322	0.050
Roll_r	0.585	0.738	0.340	0.452
Kyle	0.346	0.050	0.267	0.036
CS	0.224	0.296	0.161	0.182
AR	0.231	0.436	0.149	0.319
daily	ES	PI	ES	PI
Amihud	0.367	0.059	0.306	0.042
Roll_r	5.607	5.816	3.533	3.639
Kyle	0.308	0.057	0.242	0.049
CS	2.511	2.783	1.872	2.135
AR	2.801	3.063	2.041	2.304
15-daily	ES	PI	ES	PI
Amihud	0.358	0.054	0.320	0.435
Roll_r	31.23	31.44	21.71	21.81
Kyle	0.284	0.074	0.244	0.067
CS	15.46	15.74	13.49	13.77
AR	15.36	15.64	13.23	13.51

as benchmark measure the Kyle and Obizhaeva (2016) estimator (which performs rather poorly for the other benchmark measures) performs best. As for the pair BTCUSD the volume proxy measures are surprisingly highly positively correlated with the benchmark measures, particularly with the effective spread.

The quantile dependence plots for ETHUSD (see Figure 13) show that, besides the Corwin and Schultz (2012) and the Abdi and Ranaldo (2017) estimators, also the Kyle and Obizhaeva (2016) estimator performs well. For two of the benchmark measures, and only for low values of q, the Amihud (2002) illiquidity ratio also performs well. When we construct a composite estimator from the eight proxy measures by means of a principal component analysis we find the time-series correlations between the composite estimator and our benchmark measures to be almost universally lower than the correlations between the benchmarks and the Corwin and Schultz (2012) and the Abdi and Ranaldo (2017) estimators, respectively (see Table 11).

We separately calculate the time-series correlations between the proxy liquidity measures and the benchmark measures for the first and the second half of the sample period, for high and low return, high and low volatility, and high and low volume periods for the daily time frame. Results are reported in Figs. 14–17. The time-series correlations in the first and the second half of the sample periods are roughly similar, especially when considering *ES*, *PI* and *CRT*. Thus, and in contrast to the results for BTCUSD, we do not find markedly lower correlations in the second half of the sample period.

Correlations tend to be higher in high volume and high volatility periods for all our benchmark measures while high and low return intervals yield roughly similar correlations. Most importantly we find that the rankings of the proxy measures remain largely unchanged in the sub-samples that result from our sample splits. The Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators have the highest correlations with almost all benchmark measures under almost all conditions. When the cost of a roundtrip trade is used as the benchmark the Kyle and Obizhaeva (2016) estimator in some cases achieves better results than the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators.

To summarize, our findings relating to time-series correlations between proxy and benchmark measures for the pair ETHUSD are similar to the results for BTCUSD.

In a next step we investigate the ability of the low-frequency measures to capture the level of the high-frequency benchmarks. We use the same performance metrics as in Section 3.5 and obtain results for the pair ETHUSD that are again qualitatively similar to those for BTCUSD (see Table 12), with one important exception. The Amihud (2002) illiquidity ratio performs poorly at the hourly frequency because there are several one-hour intervals with very little volume but considerable price changes. The Kyle and Obizhaeva (2016) estimator performs best overall. The illiquidity ratio does well at lower data frequencies, particularly when the price impact is used as benchmark. As for the currency pair BTCUSD we find the Roll (1984), the Abdi and Ranaldo (2017) and the Corwin and Schultz (2012) estimators to perform poorly at lower data frequencies. Again, the performance of these proxy measures gets worse the longer the time frame, probably because of the small sample bias mentioned on page 31.

Finally, we analyze the ability of the low-frequency measures to replicate the ranking produced by the benchmark measures as in Section 3.6 above. Results are displayed in Table 13. Two measures stand out, the Amihud (2002) illiquidity ratio and the Kyle and Obizhaeva (2016) estimator. Both measures achieve a rate of correct rankings significantly above 50% in 25 out of 36 cases. The Corwin and Schultz (2012) estimator that performed well when applied to BTCUSD, does poorly when applied to ETHUSD.

4. Conclusion

In this paper we compare the performance of transactions-based liquidity measures to benchmark measures derived from high-frequency order book data. We use data for the two most actively traded cryptocurrencies, bitcoin and ethereum, and from three trading venues. We consider four benchmark measures, (a) the quoted and (b) the effective spread, (c) the price impact, and (d) the cost of a roundtrip trade, and we consider the performance of the transactions-based measures across three dimensions, (i) their ability to capture the time-series variation in liquidity, (ii) their ability to capture the level of liquidity, and (iii) their ability to capture cross-exchange differences in liquidity.

We find that no estimator performs well across all dimensions. The Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators best capture the time series variation in liquidity. This is true overall, at different quantiles of the distribution, in the first and the second half of the sample period, and in sub-samples of high and low return, high and low volatility, and high and low volume periods. The measures that perform best in the cross-sectional analysis are the Amihud (2002) illiquidity ratio and the Kyle and Obizhaeva (2016) estimator because they do well at all data frequencies and for both currency pairs. When estimating the level of the benchmark measures again the Amihud (2002) illiquidity ratio and the Kyle and Obizhaeva (2016) estimator perform best while the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) estimators do poorly in this respect.

Overall our results suggest that investors should use the Amihud (2002) illiquidity ratio or the Kyle and Obizhaeva (2016) estimator to identify the most liquid exchange. The same recommendation holds for investors when estimating the level of execution costs in order to incorporate these into their trading strategies. On the other hand, researchers looking for a measure that captures the time-series variation of liquidity, or investors hoping to time the liquidity of cryptocurrency markets and enter or exit when markets are liquid are best served by the Corwin and Schultz (2012) and Abdi and Ranaldo (2017) liquidity measures because these estimators best capture time-series

Table 8
This table reports the percentage of matching orders of proxy and benchmark liquidity measures for paired exchange-wise comparisons for the pair BTCUSD. Panels depict results for hourly, daily and 15-daily intervals respectively. Each panel reports results for the three respective pairs and the four benchmark measures. For instance, 96.84 in the hourly panel, for the pair Bitfinex / Bitstamp (BF / BS), for the quoted spread (QS) and the Abdi and Ranaldo estimator (AR) implies that the Abdi and Ranaldo estimator correctly matches the ranking of the quoted spread for this exchange pair in 96.84% of all intervals. An asterisk (*) indicates that the corresponding value is significantly higher (at the 1% level) than 50% (the value that would obtain from guessing). The sample period is 12/16/2017 to 12/16/2019.

Hourly	BF / BS	TX	$Vol	Amihud	Roll_r	Roll_p	Kyle	CS	AR
	QS	5.32	24.19	91.78*	75.71*	75.65*	89.84*	92.64*	96.84*
	ES	4.35	19.55	76.06*	61.60*	61.56*	74.35*	78.01*	80.79*
	PI	32.46	37.11	50.13	43.92	43.91	49.80	53.04*	52.57*
	CRT	5.64	24.16	91.70*	75.54*	75.48*	89.84*	92.47*	96.60*
	BF / CB								
	QS	55.80*	77.34*	46.97	38.61	38.60	30.12	69.26*	72.34*
	ES	53.52*	60.60*	45.48	37.86	37.79	39.60	67.57*	69.63*
	PI	38.60	50.71	57.81*	35.49	35.52	47.69	48.80	54.04*
	CRT	49.39	44.65	63.25*	37.10	37.13	59.65*	57.16*	58.93*
	BS / CB								
	QS	4.03	32.90	97.32*	76.70*	76.70*	86.83*	87.86*	96.28*
	ES	4.09	29.24	84.31*	65.92*	65.92*	74.77*	78.46*	84.17*
	PI	24.96	37.60	63.18*	53.23*	53.23*	59.14*	58.98*	62.81*
	CRT	4.07	32.92	97.28*	76.67*	76.67*	86.86*	87.87*	96.27*
daily	BF / BS								
	QS	0.30	20.36	81.00*	37.86	37.56	79.03*	75.72*	52.34
	ES	0.45	20.06	78.43*	36.50	36.20	76.47*	72.85*	50.98
	PI	21.42	35.29	61.09*	34.69	34.69	60.03*	65.16*	47.51
	CRT	0.30	20.36	81.00*	37.86	37.56	79.03*	75.72*	52.34
	BF / CB								
	QS	56.22*	85.16*	21.89	42.43	42.88	17.99	75.11*	61.62*
	ES	59.07*	74.96*	30.73	41.98	41.98	26.24	68.97*	58.92*
	PI	30.43	54.57	45.58	39.88	40.03	45.88	57.87*	54.42
	CRT	44.23	44.38	59.07*	42.28	42.43	57.87*	53.52	52.62
	BS / CB								
	QS	0.30	17.07	84.59*	42.90	42.45	83.08*	91.39*	61.18*
	ES	0.45	16.92	82.33*	41.24	40.79	80.21*	88.82*	59.82*
	PI	16.16	24.47	73.11*	40.33	39.88	72.51*	75.98*	56.95*
	CRT	0.30	17.07	84.59*	42.90	42.45	83.08*	91.39*	61.18*
15-daily	BF / BS								
	QS	0.00	19.57	82.61*	32.61	28.26	78.26*	58.70	58.70
	ES	0.00	19.57	82.61*	32.61	28.26	78.26*	58.70	58.70
	PI	13.04	28.26	69.57*	32.61	32.61	69.57*	58.70	58.70
	CRT	0.00	19.57	82.61*	32.61	28.26	78.26*	58.70	58.70
	BF / CB								
	QS	65.22	93.48*	8.70	39.13	43.48	15.22	78.26*	60.87
	ES	69.57*	84.78*	17.39	45.65	50.00	23.91	73.91*	60.87
	PI	17.39	50.00	52.17	39.13	36.96	50.00	56.52	43.48
	CRT	41.30	43.48	54.35	41.30	41.30	56.52	50.00	41.30
	BS / CB								
	QS	0.00	13.04	86.96*	30.43	26.09	89.13*	84.78*	47.83
	ES	0.00	13.04	86.96*	30.43	26.09	89.13*	84.78*	47.83
	PI	10.87	15.22	84.78*	23.91	19.57	86.96*	82.61*	50.00
	CRT	0.00	13.04	86.96*	30.43	26.09	89.13*	84.78*	47.83

variation in liquidity. These differing findings suggest that the the setting is important in determining the best liquidity proxy.

Our results can be used by researchers, investors, traders, and regulators to understand liquidity levels and dynamics with relatively easy to acquire and process aggregate price and volume data. In many applications, the transactions-based aggregate measures perform adequately when describing high-frequency measures derived from order book data. The use of these low-frequency measures is far less time-consuming and memory-intensive, offering a reasonable compromise between accuracy and computational workload. Strategies that require more granular data such as triangular arbitrage or market-making will of course require higher frequency measures.

CRediT authorship contribution statement

Alexander Brauneis: Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Roland Mestel:** Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Ryan Riordan:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **Erik Theissen:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing.

Acknowledgements

We thank Yakov Amihud and Hans Manner for valuable comments. We gratefully acknowledge financial support from the University of Graz.

Appendix

This appendix reports the results for the currency pair ETHUSD.

Table 9
Descriptive data for benchmark liquidity measures for the pair ETHUSD used in the empirical analysis at a daily resolution. The table reports descriptive statistics for the quoted spread (QS), the effective spread (ES), the price impact (PI) and the percentage cost of a roundtrip trade (CRT). The unit of measurement is basis points. The sample period is 12/16/2017 to 12/16/2019.

exchange		mean	std. dev.	Q1	median	Q3	num daily obs.
Bitfinex	QS	1.367	1.433	0.646	1.020	1.420	699
	ES	1.929	1.622	0.901	1.538	2.320	
	PI	0.578	0.440	0.274	0.461	0.736	
	CRT	7.224	3.809	3.854	7.634	9.885	
Bitstamp	QS	13.17	4.681	10.11	12.04	15.31	651
	ES	13.45	4.806	10.32	12.39	15.43	
	PI	0.770	0.584	0.411	0.624	0.929	
	CRT	30.33	9.934	23.17	29.15	35.24	
Coinbase	QS	1.232	0.928	0.647	0.979	1.475	676
	ES	2.177	2.294	1.157	1.790	2.580	
	PI	0.601	0.587	0.227	0.458	0.769	
	CRT	7.447	3.125	5.351	6.916	8.988	

Table 10
Descriptive data for proxy liquidity measures for the pair ETHUSD used in the empirical analysis at a daily resolution. The table reports descriptive statistics for the number of transactions (TX), the dollar volume ($ Vol, million USD), the Amihud measure (Amihud, values *1e6), Roll's returns based measure (Roll_r, basispoints), Roll's price based measure (Roll_p), the Kyle and Obizhaeva measure (Kyle, values *1e3), the Corwin and Schultz measure (CS, basispoints) and the Abdi and Ranaldo measure (AR, basispoints). The sample period is 12/16/2017 to 12/16/2019.

exchange		mean	std. dev.	Q1	median	Q3	no. daily obs
Bitfinex	TX	33,985	25,895	13,948	27,570	47,358	699
	$ Vol	59.24	72.87	13.88	33.20	77.84	
	Amihud	0.044	0.436	0.003	0.006	0.016	
	Roll_r	48.51	56.57	0	37.30	70.78	
	Roll_p	2.006	3.882	0	0.725	1.906	
	Kyle	0.131	0.041	0.101	0.124	0.156	
	CS	26.82	21.74	13.16	20.58	34.00	
	AR	30.32	24.20	15.60	23.83	36.51	
Bitstamp	TX	7,889	8,127	2,947	5,451	10,075	651
	$ Vol	12.47	17.26	3.633	7.136	14.21	
	Amihud	0.286	1.157	0.019	0.041	0.110	
	Roll_r	49.01	54.73	0	36.74	69.82	
	Roll_p	2.089	3.957	0	0.747	2.046	
	Kyle	0.212	0.060	0.171	0.203	0.242	
	CS	28.46	22.14	14.63	22.35	35.32	
	AR	31.13	23.34	16.13	24.50	38.08	
Coinbase Pro	TX	35,894	31,307	14,997	26,626	46,208	676
	$ Vol	41.60	63.87	10.11	21.31	45.18	
	Amihud	0.568	10.38	0.005	0.009	0.017	
	Roll_r	46.42	54.08	0	33.90	67.08	
	Roll_p	1.919	3.715	0	0.659	1.826	
	Kyle	0.143	0.040	0.116	0.137	0.163	
	CS	26.10	20.54	13.13	20.65	32.34	
	AR	29.80	23.62	15.07	23.43	36.01	

Table 11
This table reports mean correlations among the first principal component of proxy liquidity measures and the four benchmark measures for the pair ETHUSD. Column 'expl. var' shows the percentage of total variance explained by the first principal component. Rows denoted 'mean' contain equally-weighted averages across exchanges. The sample period is 12/16/2017 to 12/16/2019.

		expl. var	QS	ES	PI	CRT
hourly	Bitfinex	50.98	0.696	0.660	0.752	0.414
	Bitstamp	50.23	0.636	0.687	0.692	0.548
	Coinbase Pro	51.80	0.533	0.438	0.836	0.359
	mean	51.00	0.621	0.595	0.760	0.440
daily	Bitfinex	52.35	0.679	0.626	0.820	0.231
	Bitstamp	54.05	0.660	0.719	0.823	0.536
	Coinbase Pro	52.24	0.345	0.457	0.838	0.177
	mean	52.88	0.561	0.600	0.827	0.315
15-daily	Bitfinex	56.29	0.475	0.295	0.725	−0.191
	Bitstamp	50.21	0.715	0.740	0.783	0.493
	Coinbase Pro	54.80	−0.032	0.137	0.851	−0.293
	mean	53.76	0.386	0.391	0.786	0.003

Table 12
This table reports average root mean squared errors (RMSE) and average mean absolute errors (MAE) for the three exchanges for the pair ETHUSD. The panels in the table refer to hourly, daily and 15-daily results respectively. Lines depict proxy measures whilst columns in the table refer to the benchmark liquidity measures effective spread (ES) and price impact (PI). Note that Amihud is conceptually a proxy for the price impact, while the other low-frequency measures Roll_r, Kyle, CS and AR estimate the spread. All values are multiplied by 1,000. The sample period is 12/16/2017 to 12/16/2019.

	RMSE		MAE	
Hourly	ES	PI	ES	PI
Amihud	63.14	63.05	2.013	1.447
Roll_r	0.855	1.103	0.544	0.715
Kyle	0.693	0.085	0.549	0.062
CS	0.578	0.339	0.460	0.173
AR	0.485	0.584	0.347	0.424
daily	ES	PI	ES	PI
Amihud	0.655	0.085	0.575	0.065
Roll_r	6.873	7.246	4.556	4.774
Kyle	0.532	0.117	0.431	0.103
CS	2.914	3.380	2.159	2.647
AR	3.311	3.778	2.479	2.977
15-daily	ES	PI	ES	PI
Amihud	0.641	0,077	0.599	0.068
Roll_r	41.49	41.93	32.69	33.00
Kyle	0.482	0.147	0.424	0.134
CS	18.84	19.35	17.14	17.67
AR	19.08	19.58	17.06	17.60

Table 13
This table reports the percentage of matching orders of proxy and benchmark liquidity measures for paired exchange-wise comparisons for the pair ETHUSD. Panels depict results for hourly, daily and 15-daily intervals respectively. Each panel reports results for the three respective pairs for Bitfinex (BF), Bitstamp (BS) and Coinbase Pro (CB). For instance, 84.62 in the hourly panel, for the pair Bitfinex / Bitstamp (BF / BS), for the quoted spread (QS) and the Abdi and Ranaldo estimator (AR) implies that the Abdi and Ranaldo estimator matches the order of the contemporaneous order of the quoted spread for this exchange pair in 84.62% of all intervals. An asterisk (*) indicates that the corresponding value is significantly higher (at the 1% level) than 50% (the value that would obtain from guessing). The sample period is 12/16/2017 to 12/16/2019.

Hourly	BF / BS	TX	$Vol	Amihud	Roll_r	Roll_p	Kyle	CS	AR
	QS	0.34	4.99	93.75*	68.41*	68.21*	97.02*	57.29*	84.62*
	ES	0.16	1.35	38.91	26.64	26.54	39.76	27.28	36.55
	PI	11.65	12.54	28.02	20.99	21.03	28.34	25.98	27.95
	CRT	0.37	4.97	93.73*	68.39*	68.18*	97.05*	57.27*	84.60*
	BF / CB								
	QS	51.52*	50.98	54.76*	40.24	40.26	54.25*	68.28*	68.81*
	ES	34.28	28.49	44.56	29.06	29.01	45.87	52.98*	53.45*
	PI	38.25	35.73	39.61	28.23	28.16	39.03	41.98	42.90
	CRT	46.87	31.08	69.00*	41.22	41.23	71.84*	65.22*	66.29*
	BS / CB								
	QS	0.55	6.30	97.27*	66.39*	66.44*	96.98*	47.54	76.10*
	ES	0.36	1.91	40.90	26.53	26.53	40.70	24.17	34.77
	PI	16.05	16.87	25.28	20.37	20.37	25.22	23.01	24.19
	CRT	0.55	6.30	97.27*	66.39*	66.44*	96.98*	47.54	76.10*
daily	BF / BS								
	QS	0.00	0.31	97.85*	37.23	36.92	99.69*	54.00	53.08
	ES	0.00	0.31	87.54*	31.69	31.54	88.77*	49.69	47.08
	PI	20.46	20.77	67.69*	32.77	33.38	68.31*	46.77	47.85
	CRT	0.00	0.31	97.85*	37.23	36.92	99.69*	54.00	53.08
	BF / CB								
	QS	47.02	48.81	52.08	37.50	36.90	54.17	56.70*	51.49
	ES	50.30	41.96	55.80*	34.97	35.12	56.55*	57.29*	49.11
	PI	48.36	48.66	47.47	39.73	40.18	49.11	51.49	49.70
	CRT	38.54	27.53	73.96*	35.86	36.76	75.74*	50.00	52.23
	BS / CB								
	QS	0.00	0.31	97.69*	44.31	44.77	99.69*	58.92*	54.62*
	ES	0.15	0.31	86.15*	38.92	39.38	88.15*	54.15*	48.46
	PI	19.85	20.00	66.92*	37.85	38.00	68.31*	48.46	51.38
	CRT	0.00	0.31	97.69*	44.31	44.77	99.69*	58.92*	54.62*
15-daily	BF / BS								
	QS	0.00	0.00	100.00*	39.13	39.13	100.00*	45.65	45.65
	ES	0.00	0.00	100.00*	39.13	39.13	100.00*	45.65	45.65
	PI	17.39	17.39	82.61*	30.43	30.43	82.61*	50.00	32.61
	CRT	0.00	0.00	100.00*	39.13	39.13	100.00*	45.65	45.65
	BF / CB								
	QS	43.48	50.00	50.00	43.48	32.61	54.35	43.48	50.00
	ES	52.17	45.65	54.35	43.48	36.96	58.70	33.78	54.35
	PI	58.70	43.48	56.52	52.17	52.17	52.17	45.65	39.13

(*continued on next page*)

Table 13 (*continued*)

Hourly	BF / BS	TX	$Vol	Amihud	Roll_r	Roll_p	Kyle	CS	AR
	CRT	34.78	28.26	71.74*	43.48	39.13	76.09*	39.13	45.65
	BS / CB								
	QS	0.00	0.00	100.00*	30.43	34.78	100.00*	69.57*	56.52
	ES	0.00	0.00	100*	30.43	34.78	100.00*	69.57*	56.52
	PI	21.74	21.74	78.26*	32.16	36.96	78.26*	52.17	43.48
	CRT	0.00	0.00	100.00*	30.43	34.78	100.00*	69.57*	56.52

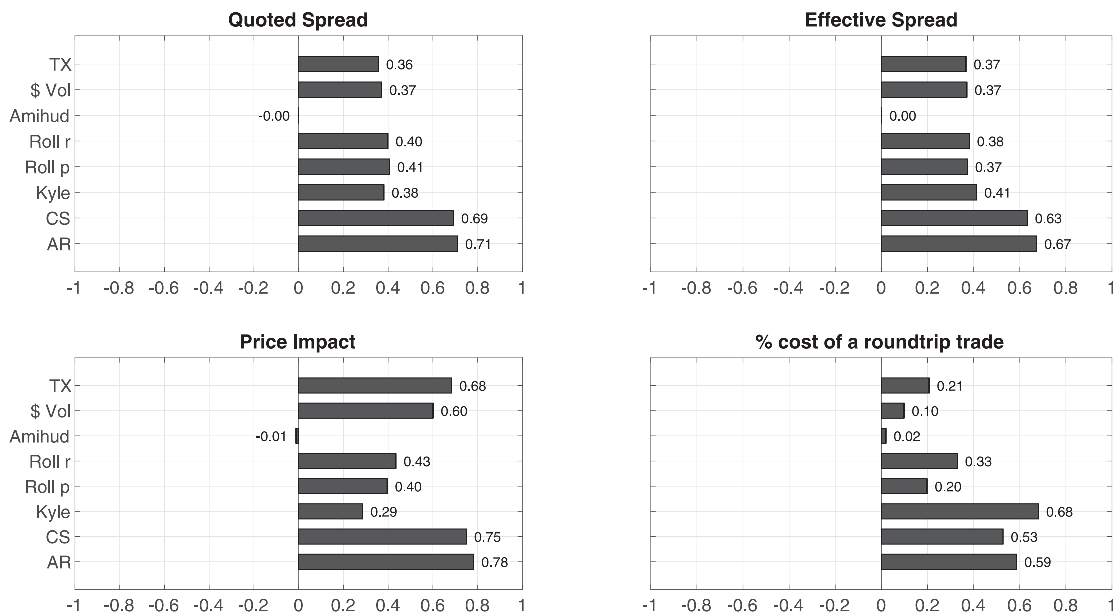

Fig. 10. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on an hourly basis over the sample period 12/16/2017 to 12/16/2019.

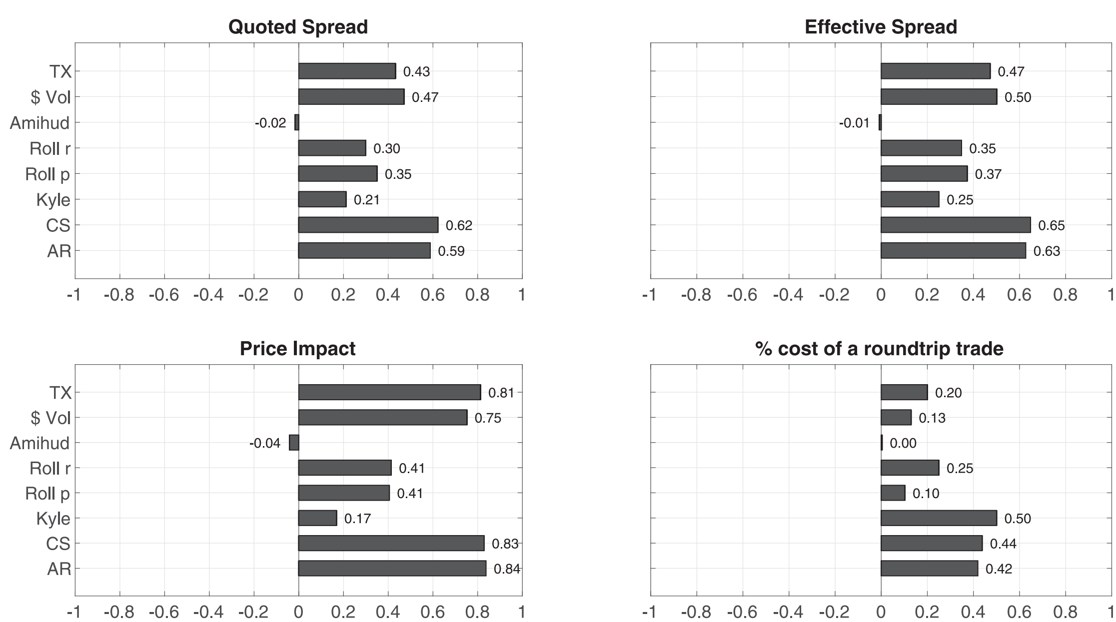

Fig. 11. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

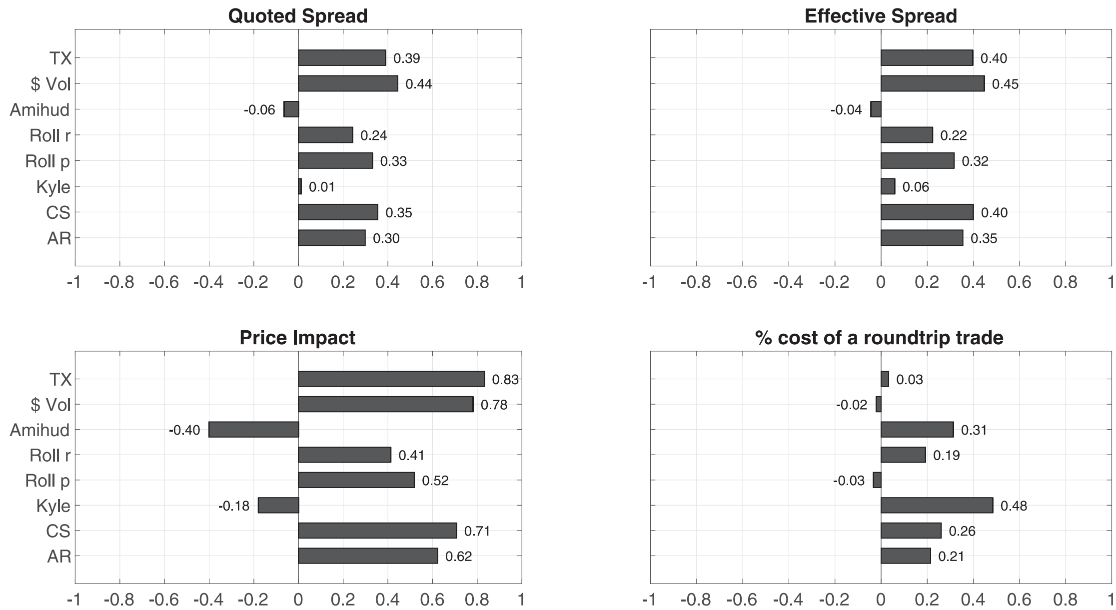

Fig. 12. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a 15-daily basis over the sample period 12/16/2017 to 12/16/2019.

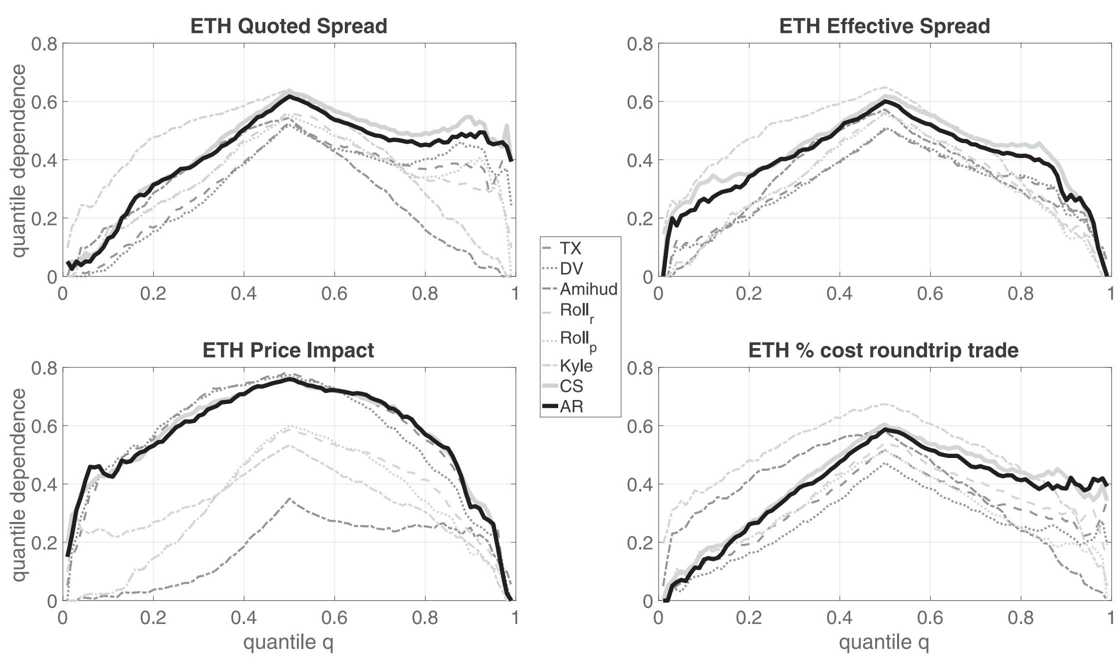

Fig. 13. This figure plots the quantile dependence (averaged across the three trading venues) as a function of the quantile q in steps of 0.01 for the four benchmark liquidity measures and each of the 8 proxy measures for the pair ETHUSD for daily data. The sample period is 12/16/2017 to 12/16/2019.

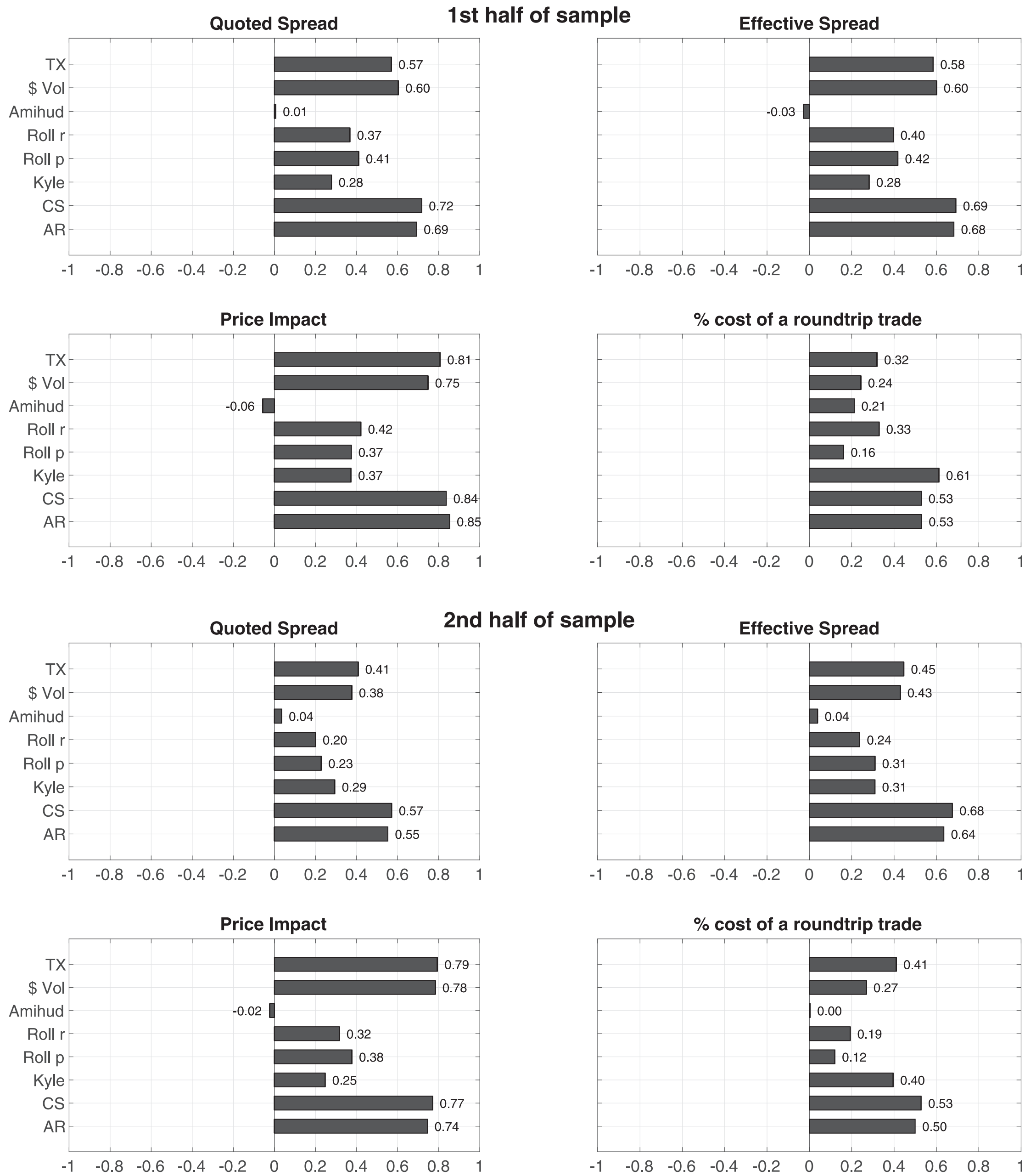

Fig. 14. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures for the first and second half of the sample for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

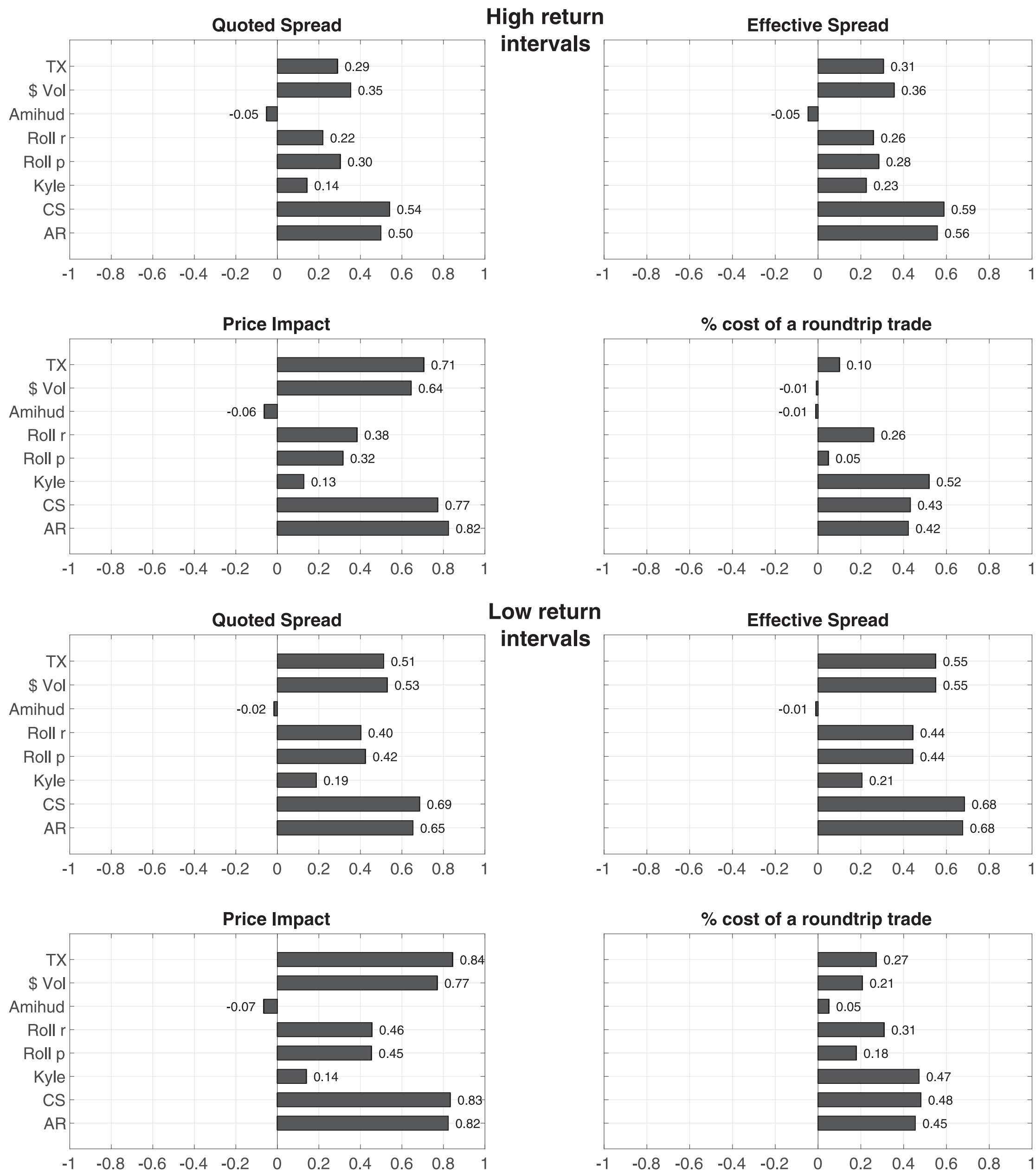

Fig. 15. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low return intervals for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

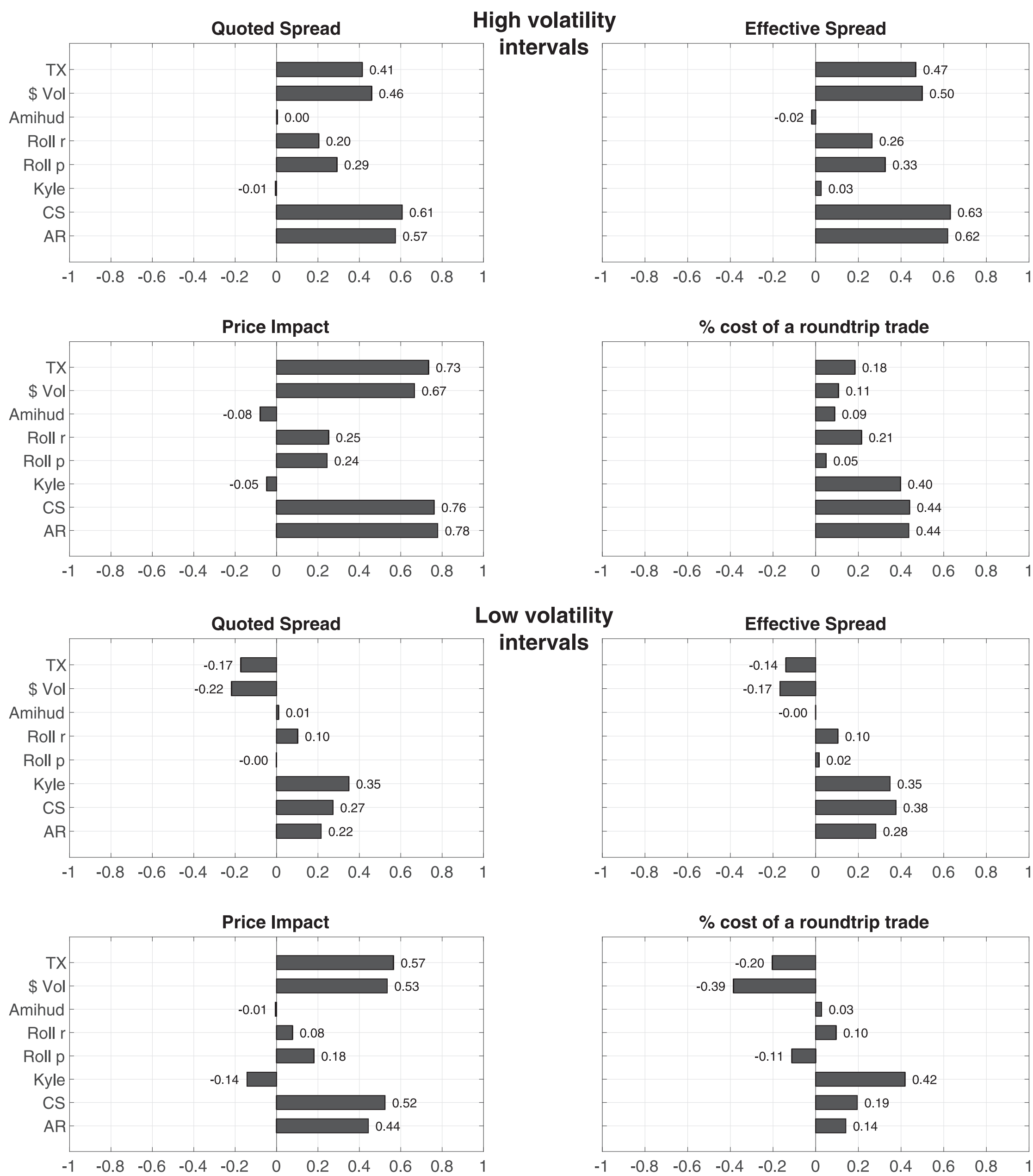

Fig. 16. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low volatility intervals for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

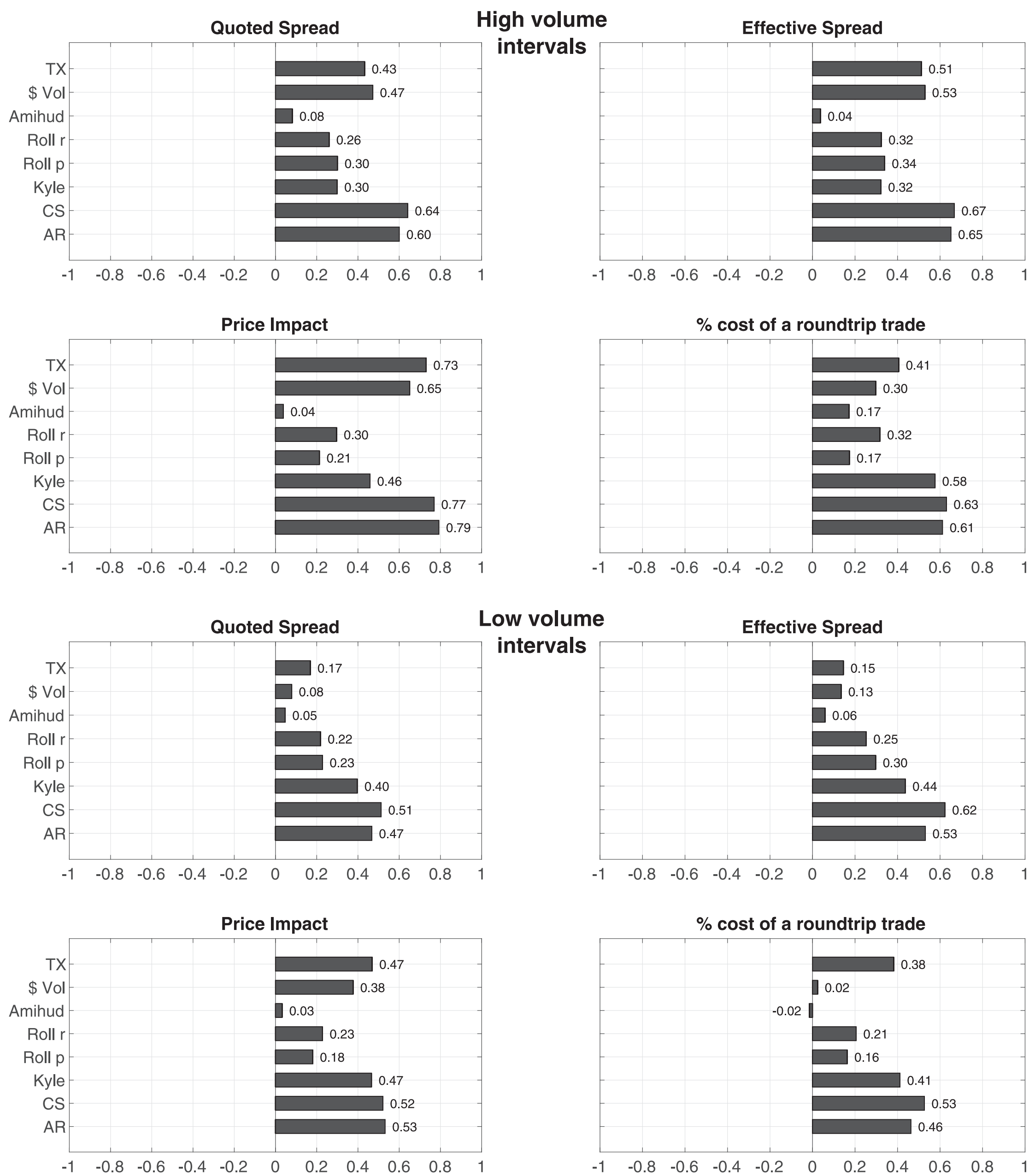

Fig. 17. This figure shows time series correlations between benchmark measures of liquidity and low-frequency proxy measures in the subset of high and low volume intervals for the pair ETHUSD. Values represent simple averages across the three exchanges Bitfinex, Bitstamp, and Coinbase Pro. Liquidity measures are calculated on a daily basis over the sample period 12/16/2017 to 12/16/2019.

References

Abdi, F., Ranaldo, A., 2017. A simple estimation of bid-ask spreads from daily close, high and low prices. Rev. Financ. Stud. 30 (12), 4437–4480.

Alexander, C., Dakos, M., 2019. A critical investigation of cryptocurrency data and analysis. Quant. Finance 20 (2), 173–188.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. J. Financ. Mark. 5 (1), 31–56.

Amihud, Y, Noh, J, 2002. Illiquidity and stock returns II: cross–section and time-series effects. Review of Financial Studies https://doi.org/10.1093/rfs/hhaa080. In press.

Bogousslavsky, V., Collin-Dufresne, P., 2020. Liquidity, volume, and volatility. Available at SSRN: https://ssrn.com/abstract=3336171.

Brauneis, A., Mestel, R., 2018. Price discovery of cryptocurrencies: bitcoin and beyond. Econ. Lett. 165, 58–61.

Brauneis, A., Mestel, R., Riordan, R., Theissen, E., 2021. Bitcoin unchained: Determinants of cryptocurrency exchange liquidity. Available at SSRN: https://ssrn.com/abstract=3249477.

Conrad, J., Wahal, S., 2020. The term structure of liquidity provision. J. Financ. Econ. 136, 239–259.

Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. J. Finance 67 (2), 719–760.

Dimpfl, T., 2017. Bitcoin Market Microstructure. Available at SSRN: https://ssrn.com/abstract=2949807.

Duan, F., Manner, H., Wied, D., 2019. Model and moment selection in factor copula models. J. Financ. Econom. 28, 1–31.

Dyhrberg, A., Foley, S., Svec, J., 2018. How investible is bitcoin? Analyzing the liquidity and transaction costs of bitcoin markets. Econ. Lett. 171, 140–143.

Fink, C., Johann, T., 2014. Bitcoin Markets. Available at SSRN: https://ssrn.com/abstract=2408396.

Fong, K.Y., Holden, C.W., Trzcinka, C.A., 2017. What are the best liquidity proxies for global research? Rev. Finance 21 (4), 1355–1401.

Goyenko, R.Y., Holden, C.W., Trzcinka, C.A., 2009. Do liquidity measures measure liquidity? J. Financ. Econ. 92 (2), 153–181.

Harris, L., 1990. Statistical properties of the roll serial covariance bid/ask spread estimator. J. Finance 45 (2), 579–590.

Hasbrouck, J., Seppi, D.J., 2001. Common factors in prices, order flows, and liquidity. J. Financ. Econ. 59 (3), 383–411.

Hautsch, N., Scheuch, C., Voigt, S., 2018. Building Trust Takes Time: Limits to Arbitrage in Blockchain-Based Markets. Available at SSRN: https://ssrn.com/abstract=3302159.

Hougan, M., Kim, H., Lerner, M., 2019. Economic and non-economic trading in bitcoin: exploring the real spot market for the world's first digital commodity. Working Paper. Available at: https://www.sec.gov/comments/sr-nysearca-2019-01/srnysearca201901-5574233-185408.pdf

Hu, A.S., Parlour, C.A., Rajan, U., 2019. Cryptocurrencies: stylized facts on a new investible instrument. Financ. Manag. 48 (4), 1049–1068.

Irvine, N., Benston, G., Kandel, E., 2000. Liquidity Beyond the Inside Spread: Measuring and Using Information in the Limit Order Book. Available at SSRN: https://ssrn.com/abstract=229959.

Johann, T., Theissen, E., 2020. The Best in Town: A Comparative Analysis of Low-Frequency Liquidity Estimators. Available at SSRN: https://ssrn.com/abstract=2905032.

Karnaukh, N., Ranaldo, A., Söderlind, P., 2015. Understanding FX liquidity. Rev. Financ. Stud. 28, 3073–3108.

Kyle, A.S., Obizhaeva, A.S., 2016. Market microstructure invariance: empirical hypotheses. Econometrica 84, 1345–1404.

Liu, Y., Tsyvinski, A., Wu, X., 2019. Common risk factors in cryptocurrency. NBER Working Paper 25882. Available at: https://www.nber.org/papers/w25882

Makarov, I., Schoar, A., 2020. Trading and arbitrage in cryptocurrency markets. J. Financ. Econ. 135 (2), 293–319.

Mancini, L., Ranaldo, A., Wrampelmeyer, J., 2013. Liquidity in the foreign exchange market: measurement, commonality, and risk premiums. J. Finance LXVIII, 1805–1841.

Marshall, B., Nguyen, N., Visaltanachoti, N., 2019. Bitcoin Liquidity. Available at SSRN: https://ssrn.com/abstract=3194869.

Marshall, B.R., Nguyen, N.H., Visaltanachoti, N., 2012. Commodity liquidity measurement and transaction costs. Rev. Financ. Stud. 25 (2), 599–638.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. J. Finance 39 (4), 1127–1139.

Schestag, R., Schuster, P., Uhrig-Homburg, M., 2016. Measuring liquidity in bond markets. Rev. Financ. Stud. 29 (5), 1170–1219.

Shi, S., 2018. The Impact of Futures Trading on Intraday Spot Volatility and Liquidity: Evidence from Bitcoin Market. Available at SSRN: https://ssrn.com/abstract=3094647.