LUISS

Department
of Economics and Finance

Course of Asset Pricing

Liqudity in cryptocurrency market

Prof. Borri

SUPERVISOR

Prof. Carlini

CO-SUPERVISOR

723971

CANDIDATE

Academic Year 2019/2020

Contents

1. Introduction

Nowadays, in the modern financial markets with a wide variety of different investment opportunities, liquidity became a vital decision indicator. The impact of liquidity on the choice of investment draws more and more attention. Liquidity is often defined as "an ability to trade large quantities quickly at low cost with little price impact."

Cryptocurrencies have less history than any other traditional asset, but they draw significant attention from investors and researches. Since the introduction of bitcoin in 2009, they have become increasingly popular. Cryptocurrency is a digital token that can be exchanged online. It uses cryptographic hashing and digital signatures to verify transactions and avoid double-spending of the same token. These features prevent from creating tokens from the air, so the supply is limited. In this aspect, cryptocurrencies are like metals or other commodities. Because the limited quantity of cryptocurrencies is protected by the cryptography embedded in their open-source code, cryptocurrencies can potentially become valuable.

There is a wide range of papers written on the topic of liquidity. They investigate liquidity measures, construct liquidity factors, and measure their impact on asset prices. Cryptocurrencies also raise interest, so despite the relative youthfulness, there are many papers on this topic. However, there are almost no investigation on the topic of liquidity in the crypto market and its connection with different factors, including traditional markets. Cryptocurrencies have some peculiar properties that can prevent us from using traditional techniques for liquidity estimating.

The main aim of this paper is to fill this gap. Therefore, the objectives of this research are the following:

- Determine the range of existing liquidity proxies in the absence of quote data
- Validate proxies
- Explore factors affecting the liquidity of the cryptocurrency market
- Investigate how liquidity affects portfolio optimization and rebalancing strategy

This investigation stands out among others, due to the following reasons. First, we explore and compare different liquidity estimators based on low-frequency data. It allows us to use more available data sources without restrictions. Second, we check the performance of the estimators in the cryptocurrency markets and point out the limitations of existing estimators for the crypto market. Third, we run a comprehensive investigation of different factors affecting cryptocurrencies' liquidity. Finally, we explore how transaction costs may affect different portfolios' performance.

This investigation is divided into the following sections: literature review, data description, and data analysis, review of different coins and investigation of the applicability of liquidity measures, cross-sectional analysis of liquidity factors, and time-series analysis of liquidity factors using both separate coins analysis and panel analysis, portfolio optimization, conclusion.

2. Literature review

Liquidity is a concept that is hard to define and hard to measure. That is why researches started inventing some proxies to measure liquidity. One of these proxies is bid-ask spread, but to explicitly measure quoted spread, we need to have tick data, which is usually expensive, hard to collect, and process. That is why, in 1984, Richard Roll suggested a measure based only on price data (Roll 1984). He proved that effective bid-ask spread could be measured by

$$Spread = 2\sqrt{-cov}$$

where “cov” is the first-order serial covariance of price changes. Roll verified his estimator by relating measured spread to firm size. Firm size has a positive correlation to volume, and volume has a negative correlation with spread (Demsetz 1968; Copeland and Galai 1983). Therefore, it means that there has to be a strong negative correlation between size and spread in cross-sectional regression. We will use this method for verification of estimates for cryptocurrencies, but we will use volume instead of size. It allows us to increase accuracy because we measure volume directly and not through size.

In 1988 J. Y. Choi, Dan Salandro, and Kuldeep Shastri pointed out that Roll measure does not take into account serial autocorrelation of transaction type (Choi, Salandro, and Shastri 1988). They extended Roll model and got

$$Spread = \sqrt{-cov}/(1-\delta)$$

where δ is the conditional probability that the transaction at time t-1 is at the bid (ask) price, given that the transaction at time t is at the bid (ask) price. If δ > 0.5, then roll estimator is downward biased, but this bias decrease with longer time ranges.

Later Lesmond, Ogden, and Trzcinka (Lesmond, Ogden, and Trzcinka 1999) and Hasbrouck (Hasbrouck 2004) developed additional proxies for the effective spread using daily data. Amihud (Amihud 2002) and Pastor and Stambaugh (Pástor and Stambaugh 2003) developed low-frequency liquidity measures that perhaps might be viewed as proxies for price impact, more than for the effective spread. Amihud’s work arouses much interest. Several papers replicate and extend it (Brennan, Huh, and Subrahmanyam 2013; Harris and Amato 2019). In 2009 Holden provided several other measures (Holden 2009), and Goyenko et al. proved their efficiency comparing to previous ones (Goyenko, Holden, and Trzcinka 2009).

In 2012 Corwin and Schultz developed a new estimator based on an interesting insight. They say high-low ration reflects both the real variance and bid-ask spread. Bid-ask spread is constant over time, while variance increase with a more extended period. We can solve for both the spread and the variance by deriving two equations, the first a function of the high-low ratios on two consecutive days and the second a function of the high-low ratio from a single 2-day period (Corwin and Schultz 2012).

In 2014 Chung and Zhang examined the relation between bid-ask spread calculated based on daily data of closing prices and bid-ask spread calculated based on intra-day data. It shows a very high correlation between them (Chung and Zhang 2014). We use his CRSP Spread as a benchmark for others as we do not have intraday data available.

One of the latest estimators was developed in 2017 by Abdi and Ronaldo. It uses daily high, low, and close price, extending and outperforming Roll, and other previously mentioned estimators (Abdi and Ranaldo 2017). We can calculate monthly-corrected spread by

$$Spread = \sqrt{\max\left\{4\frac{1}{N}\sum_{t=1}^{N}(c_t - \eta_t)(c_t - \eta_{t+1}),0\right\}}$$

and two-day corrected by

$$Spread = \frac{1}{N}\sum_{t=1}^{N} s_t\,, where\ s_t = \sqrt{\max\{4(c_t - \eta_t)(c_t - \eta_{t+1}),0\}}$$

where N shows a number of points in the period, c_t shows the closing price at time t and η_t shows mid-price at time t, which is equal to $\frac{h_t+l_t}{2}$, average between high and low prices. All prices are in logs.

Several papers prove the connection between liquidity and stock returns. In 1986 Amihud and Mendelson found that asset returns are increasing function of bid-ask spread (Amihud and Mendelson 1986). In 1997 Eleswarapu confirmed this model using Nasdaq data (Eleswarapu 1997).

Chalmers and Kadlec in 1998 examined amortized spreads for stocks over the period 1983-1992. The amortized spread measures the spread's cost over investors' holding periods and is approximately equal to the spread times share turnover. They find more persuasive evidence that amortized spread is priced, than for quoted spread.

In 2005, Acharya and Pedersen built a model for understanding the various channels through which liquidity risk may affect asset prices. They find a persistent negative shock to a security's liquidity results in low contemporaneous returns and high predicted future returns (Acharya and Pedersen 2005).

Another significant group of papers is devoted to finding common liquidity factors. Almost at the same time, Hasbrouck and Seppi (Hasbrouck and Seppi 2001), Huberman and Halka (Huberman and Halka 2001), Chordia and Roll (Chordia, Roll, and

Subrahmanyam 2000) published papers on this topic. They find empirical evidence of the commonality of liquidity among different assets. Later Brockman et al. find evidence of commonality within stock exchanges and across exchanges (Brockman, Chung, and Prignon 2009).

In 2007 Cheng explored factors that may influence stock liquidity (Cheng 2007). They are firm size, the centralization of the ownership structure, the degree of information asymmetry, the level of margin trading utilization, the absorbed stocks of investors, and the liquidity of the entire market. We can use some of these factors in the cryptocurrency market.

The cryptocurrency market is different from the regular stock market, and some models that work in the stock market may not work in the crypto world. Cryptocurrencies are still relatively young, but they bring the attention of the researches nowadays.

Sha Wang and Jean-Philippe Vergne, in their paper (Wang and Vergne 2017), discussed factors that affect cryptocurrency returns. They concluded that public interest negatively affects returns, while liquidity shows a positive effect. Unexpected supply also positively affects returns, which is a different behavior from fiat currencies. They provide two mechanisms that can explain it. First, an increase in supply can motivate holders to reinforce positions, and it can bring the attention of outsiders without prior awareness to participate and buy coins. Second, increased supply is most likely the result of an increase in mining intensity, and that could be interpreted as a signal of increasing potential.

Andrew Urquhart (Urquhart 2016) and Wang Chun Wei (Wei 2018) confirm that cryptocurrency market efficiency correlates with liquidity. They explain it by boom-bust speculative cycles for low liquid altcoins. Another significant result is that there is no illiquidity premium for investors in cryptocurrencies. It is a contrary result to traditional assets and currencies.

In the NBER working paper, Yukun Liu and Aleh Tsyvinski conducted very informative research on factors for cryptocurrency returns (Liu et al. 2018). They find that the CAPM model works but has a strongly significant alpha. However, cryptocurrencies have not exposure to fiat currencies, metals, and macro factors of the economy. They confirm the previously mentioned connection with investor attention and time-series momentum.

Finally, in another NBER working paper by Yukun Liu, Aleh Tsyvinski, and Xi Wu (Liu et al. 2019) and in another paper by Dehua Shen, Andrew Urquhart, and Pengfei Wang (Shen, Urquhart, and Wang 2019), researchers developed in parallel three-factor model of cryptocurrency pricing.

$$r_{i,t} - rf_t = \alpha + \beta_1 CMKT + \beta_2 CSMB + \beta_3 CMOM$$

CMKT – market excess of returns. Market return is capitalization-weighted returns of cryptocurrencies in the market.

CSMB – size factor, which defined as a return difference between portfolios of small and big sizes.

CMOM – momentum factor (Carhart 1997), which is defined as a return difference between high return portfolio and low return portfolio.

Cryptocurrency liquidity has not been widely investigated because the cryptocurrency world is relatively young. There are still many white spots in this area. It provides opportunities for further research and makes this paper valuable.

3. Data description and analysis

3.1. Data

We collect daily OHLCV data for the period from 01.01.2018 to 01.03.2020. We use the BitFinex cryptocurrency exchange with provided open API and Bloomberg as data sources. From BitFinex API, we select 20 cryptocurrencies with the highest trading volume, excluding stablecoins. Selected coins are Bitcoin, Ethereum, EOS, Ripple, Litecoin, Monero, Iota, Zcash, Dash, NEO, Ethereum Classic, Eidoo, OmiseGO, Bitcoin Gold, ETP, Streamr, Golem, Santiment, Qtum, and YOYOW.

Blomberg database has access only to nine currencies. They are Bitcoin, Dash, EOS, Ethereum, Ethereum Classic, Litecoin, Monero, Ripple, and Zcash. Besides OHLCV data, we also collect quoted bid, ask, and circulating supply of coins. The circulating supply of coins is the number of coins available for trading and equal to mined coins minus unreachable coins. If we do not need additional data on bid, ask, and supply from Bloomberg, and if we do not say it implicitly, we use BitFinex API data.

To include more liquidity factors, we collect daily data for spot gold prices, including OHLCV data and quoted bid and ask. For the same reasons, we collect VIX index historical daily data. We take all the mentioned data from Bloomberg.

We also use Google trend data, which represents a number of google searches for the topic. As the topic, we use currency name. Google normalizes data in the following way: First, it divides the number of searches for the topic on the total number of searches in the region. Second, it scales each data point on a range of 0 to 100 based on a topic’s proportion to all searches on all topics (“FAQ about Google Trends Data - Trends Help” n.d.).

$$GoogleTrend = \frac{g_{t,i}}{\max(g_{t,i})} * 100, where\ g_{t,i} = \frac{Searches\ for\ the\ topic\ i}{Total\ searches\ at\ time\ t}$$

where t – time, i – topic index. The main problem with this normalization is that google allows comparing only five topics at one time. So it becomes challenging to measure cross-sectional relations. To bypass this limitation, we use two overlapping datasets and scale them to one dimension.

For, example if we need to compare nine topics, we use two datasets with five topics, while one topic is common for both datasets. Let time series be x1, x2, x3, x4, x5 for the first dataset and y1, y2, y3, y3, y5 for the second dataset. X1 is the same topic as y1. To compare the first dataset with the second dataset, we need to calculate average x1 and average y1. Ratio $r = \frac{avg(x1)}{avg(y1)}$ is the ratio between datasets. Therefore, to compare the first and second datasets, we need to multiply the second dataset by r.

3.2. Coins review

In this section, we review and compare several major cryptocurrencies. Cryptocurrencies have one universal principle – using blockchain as a database for transactions. Unlike traditional databases, blockchain has built-in rewriting protection, so it is impossible to cancel the transaction.

Bitcoin is the first cryptocurrency and remains the most popular one. Satoshi Nakamoto created it in 2008 (Nakamoto 2008). Whet it was just created, emission of bitcoin was 50 coins every 10 minutes. Every approximately four years, there is halving in emission, so now Bitcoin has emission 6.25 coins about every 10 minutes.

Bitcoin has many forks that use the same codebase with minor or no changes. Only a few of them became famous. One of such coins is Litecoin, created in 2011. It uses the same principles as bitcoin, but the only difference is the time between blocks. Instead of 10 minutes, there are 2.5 minutes between blocks. It makes transactions faster. Block reward also halves every four years and now is equal to 12.5 coins.

Another Bitcoin fork, Dash, created in 2014, has more changes. It introduces masternodes to ensure the blockchain is readily available to all network participants and perform many other functions related to the health and efficiency of the network. It allows performing instantaneous transactions and anonymous transactions by using built-in mixers.

Zcash, created in 2014, is also based on the bitcoin protocol, but special attention is paid to privacy. It has two types of addresses: t-address works similar to bitcoin, while z-address uses zero-knowledge proof to hide transaction source, destination, and amount. However, it was found severe security issues in the anonymous part of the network leading to deanonymization (Kappos et al. 2018).

Monero, also released in 2014, is the first currency in our list, not having any connection with bitcoin. Its main feature is a complete anonymity of transactions and enforced security, so no user can accidentally be traceable or insecure. Therefore, it is widely used to hide transactions in darknet markets and to break the connection between bitcoin transactions.

Ethereum, created by Vitaly Buterin in 2014, is a platform for distributed applications, using Ether as an internal currency (Buterin 2014). It provides more flexibility than bitcoin by allowing creating complex smart contracts and lower time between blocks. It is a tipping point in the crypto world because it opens an era of decentralized applications.

In 2016, frauds managed to exploit the vulnerability in DAO project smart contract and steal $50 million worth Ether. In response to it, the community decided to rewrite the blockchain history to cancel this transaction. Part of the community ignored the attempt to rewrite history and continued to use the original network, which gets the name Ethereum Classic.

EOS.IO is another smart contract platform, created by the private company block.one in 2018. It uses different consensus algorithm proof-of-stake instead of proof-of-work. It allows users to conduct transactions faster and to reduce transaction fees. Block.one maintains the development and popularization of the platform, while blockchain is public.

Ripple is a real-time gross settlement system, currency exchange, and remittance network created by Ripple Labs Inc. in 2012. It supports tokens representing fiat currency, cryptocurrency, commodities, or other units of value. Native cryptocurrency of the system is known as XRP and positioning as SWIFT replacement. However, banks avoid using XRP currency due to its high volatility.

3.3. Prices and Returns

If we look at the price dynamics, we can see that all currencies share the same trend (Figure 1). However, they have different exposure to common shocks. For example, Bitcoin has a sharp rise in the middle of 2019, while Dash shows almost no change. It gives us a clue of having some kind of commonality among these currencies.

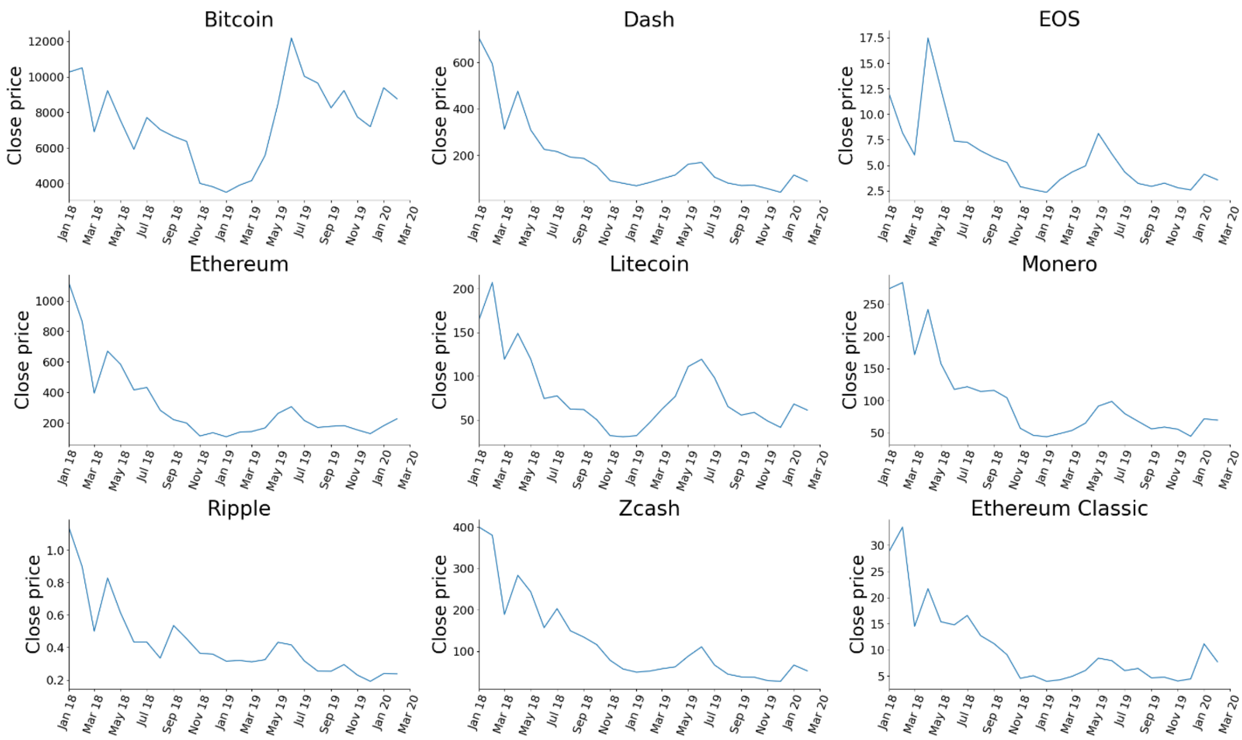

Figure 1. Close prices.

Analysis of returns (Figure 2) confirms previously mentioned commonality in factors affecting prices. However, currencies have different sensitivity to the shocks. It indirectly confirms the Capital Asset Pricing Model because different sensitivity can be explained by different betas.

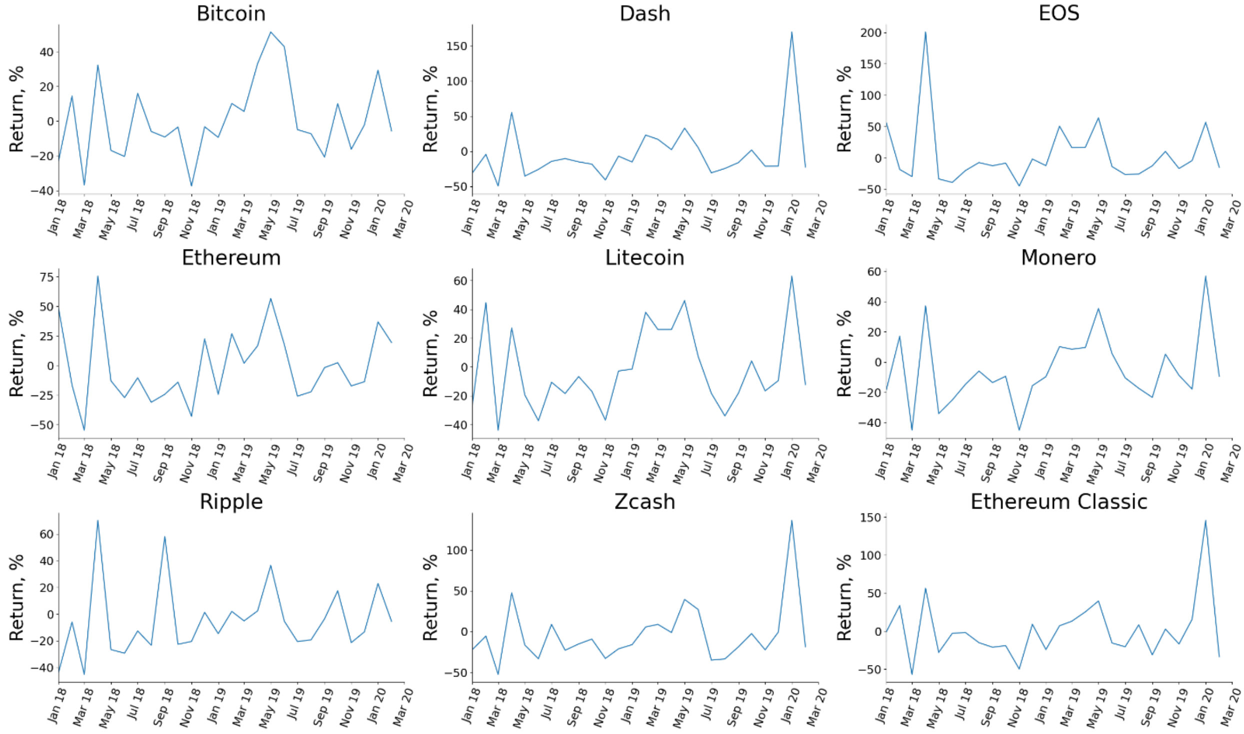

Figure 2. Returns.

3.4. Spread

Before we start working with spreads, it is essential to verify estimators. Cryptocurrencies show different behavior from traditional assets. It means that some assumptions made by estimator's authors may not hold in cryptocurrency markets. We run our checks on Roll estimator, Corwin and Schultz HL estimator, Abdi and Ronaldo CHL estimator.

Our first check is almost equal to the check conducted by Roll to check his estimator (Roll 1984). We know that spread has to have negative cross-sectional relation with volume. Several papers point to that (Demsetz 1968; Copeland and Galai 1983). Therefore, we run the following Fama-Macbeth cross-sectional relation for every estimator.

$$Spread = c + \beta_1 Volume$$

Table 1. CHL estimator regression on volume

	Value	S.D.	P-Value
Constant	2.4164	0.0998	<0.0001
Volume ($blns)	-0.9329	0.1850	<0.0001

Table 2. HL estimator regression on volume

	Value	S.D.	P-Value
Constant	2.4586	0.1405	<0.0001
Volume ($blns)	-0.4387	0.1252	0.0005

Table 3. Roll estimator regression on volume

	Value	S.D.	P-Value
Constant	3.6397	0.3032	<0.0001
Volume ($blns)	-0.9317	0.3253	0.0043

As we can see from the tables, every estimator has a significant negative relation with volume. It means these estimators have a relation to the real spread.

Our second check is measuring the relation between estimators and the benchmark of real real spread. We use the CRSP spread estimator by Chung and Zhang (Chung and Zhang 2014) as the benchmark. We check both cross-sectional (Table 5) and time-series (

Table 6) correlations.

Table 4. Spread estimators description

	Observations	Mean	S.D.	Min	Max
Roll	198	0.025855	0.028756	0	0.116382
HL	198	0.017568	0.006567	0.005131	0.041515
CHL	198	0.019741	0.009655	0.005086	0.064004
CRSP	198	0.000845	0.000674	0.000015	0.002661

From Table 4, we can see all estimators have an upward bias from the CRSP estimator. One possible explanation is negative autocorrelation in transaction type. We can make such a conclusion because we know about time-series momentum in cryptocurrencies (Liu et al. 2018). It is possible to suppose that the probability of getting the opposite transaction type is more than 50%. As we can see from the paper by Choi et all, it can lead to upward bias (Choi, Salandro, and Shastri 1988). Of course, it is just a hypothesis, and to check it, we need intraday tick data, which we do not have.

Table 5. Cross-sectional correlations

	CRSP	CHL	HL	Roll
CRSP	1	0.37830	0.27150	0.08672
CHL		1	0.62496	0.34162
HL			1	0.18337
Roll				1

Table 6. Time-series correlations

	CRSP	CHL	HL	Roll
CRSP	1	0.27463	0.30951	0.15703
CHL		1	0.60321	0.44450
HL			1	0.40424
Roll				1

For cross-sectional and time-series correlation, all estimators show much weaker performance that in the paper by Abdi and Ronaldo (Abdi and Ranaldo 2017). As we already mentioned, the cryptocurrency market is inefficient and has time-series autocorrelation. However, estimators rely on the assumption of prices moving in Geometric Brownian Motion, which is not valid in the case of cryptocurrencies. It means that the use of these estimators is very limited in cryptocurrency markets.

We focus our research on the CRSP estimator because it does not depend on assumptions needed for other estimators. We also use CHL and HL estimator as they allow us to use larger data sets. Nevertheless, in the interpretation of the results, priority goes to the CRSP estimator.

4. Factors affecting liquidity

4.1. Cross-sectional analysis

First, we examine cross-sectional factors that affect the spread. We use Google Trend value, size, or market capitalization of the currency and trading volume in dollars. All variables are in logs. We test four models by running Fama-Macbeth regressions, and then we compare "R^2 between". "R^2 between" measures the goodness of fit using the time-series average of the variables. Therefore, in our case, it measures how good our variables explain cross-sectional factors.

Following the Fama-Macbeth procedure, we run T cross-sectional regressions and then take the average of the coefficients. We use four models. Model 1 includes only Google Trend, Model 2 adds Size to Model 1, Model 3 adds Volume to Model 1, and Model 4 includes all variables:

$$Ln(Spread_i) = \beta_1 + \beta_2 * Ln(GoogleTrend_i) + \beta_3 * Ln(Size_i) + \beta_4 * Ln(Volume_i)$$

Table 7 Cross-sectional spread regression. Bloomberg data. CRSP spread.

	Model 1	Model 2	Model 3	Model 4
Ln(GoogleTrend)	-0.2813*** (0.0044)	0.0189 (0.0138)	-0.0092 (0.0057)	0.0169 (0.0108)
Ln(Size)	-	-0.5470*** (0.0284)	-	-0.2233* (0.0830)
Ln(Volume in USD)	-	-	-0.5780*** (0.0273)	-0.3684*** (0.0844)
const	-8.8455*** (0.0432)	4.7353*** (0.6406)	5.4872*** (0.5523)	5.7418*** (0.3794)
R^2	0.3138	0.7924	0.6474	0.7295
R^2 Between	0.3684	0.9309	0.9714	0.9637
R^2 Within	-0.1432	-0.3671	-2.0659	-1.2316

P-values: *** - < 0.0001, ** - < 0.001, * - < 0.01

In Table 7, we present four models of spread factors. Model 1 includes only Google Trend value and can explain only 37% of spread differences among currencies. In the other three models, Google Trend is insignificant. It means that Google Trend's influence is already included in Size or Volume. Therefore, we can check it by regressing size and volume on google trend. We got 43% and 38% values of R^2 between accordingly.

Model 2, 3, and 4 explain cross-sectional differences in spreads well enough. If we look at these models, we can notice multicollinearity in Model 4 because of the close connection between Size and Volume (VIF > 10 in model 4). All models have high values of R^2 between, so they can explain Spread very well. Therefore, we can use models 2 or 3, depending on the data available. It is a beneficial finding because it proves we can use volume or size as the controls in time series effects to capture entity effects.

We also verify this statement by checking other spread estimates and other datasets. We run the same regressions using CHL and HL spreads (Table 8 and Table 9). We also run it on a bigger dataset by BitFinex. As we have no size data in this dataset, there are only two models to check (Table 10).

Table 8 Cross-sectional spread regression. Bloomberg data. CHL spread.

	Model 1	Model 2	Model 3	Model 4
Ln(GoogleTrend)	-0.0432** 0.0107	-0.0122 (0.0207)	-0.0312** (0.0087)	-0.0121 (0.0171)
Ln(Size)	-	-0.0546 (0.0404)	-	-0.1526* (0.0588)
Ln(Volume in USD)	-	-	-0.0276 (0.0170)	0.1220* (0.0379)
const	-4.2487*** (0.1033)	-2.8800* (1.0143)	-3.5306*** (0.4341)	-3.4767** (0.8838)
R^2	0.0231	0.0249	-0.0054	0.1098
R^2 Between	0.3521	0.5358	0.3496	0.7185
R^2 Within	-0.0149	-0.0342	-0.0465	0.0394

P-values: *** - < 0.0001, ** - < 0.001, * - < 0.05

Table 9 Cross-sectional spread regression. Bloomberg data. HL spread.

	Model 1	Model 2	Model 3	Model 4
Ln(GoogleTrend)	-0.0290* (0.0139)	0.0113 (0.0089)	-0.0105* (0.0049)	0.0102 (0.0078)
Ln(Size)	-	-0.0754** (0.0182)	-	-0.2249*** (0.0150)
Ln(Volume in USD)	-	-	-0.0460 (0.0303)	0.1692*** (0.0215)
const	-4.3272*** (0.1097)	-2.4621** (0.3945)	-3.2092*** 0.6765	-2.9679*** (0.4262)
R^2	0.0176	0.0615	-0.0388	0.2635
R^2 Between	0.1760	0.6265	0.3608	0.8746
R^2 Within	-0.0154	-0.0563	-0.1221	0.1361

P-values: *** - < 0.0001, ** - < 0.001, * - < 0.05

We can see CHL and HL spreads perform worse than CRSP spread. It shows similar results, but the significance of variables may differ. These estimators are less precise; that is why we need more coins for cross-sectional regression. We use the BitFinex dataset, which includes 20 coins.

Table 10 Cross-sectional spread regression. BitFinex data.

	Model 1 (CHL)	Model 3 (CHL)	Model 1 (HL)	Model 3 (HL)
Ln(GoogleTrend)	-0.1339*** (0.0121)	-0.0321 (0.0164)	-0.1157*** (0.0081)	-0.0822*** (0.0025)
Ln(Volume in USD)	-	-0.0766*** (0.0181)	-	-0.0260*** (0.0054)
const	-4.7139*** (0.1233)	-2.8189*** (0.3352)	-4.5695*** (0.1105)	-3.9311*** (0.0465)
R^2	0.1909	0.2293	0.1372	0.1181
R^2 Between	0.5824	0.7639	0.6600	0.7025
R^2 Within	-0.1701	-0.2635	-0.1841	-0.2410

P-values: *** - < 0.0001, ** - < 0.001, * - < 0.05

As we can see from Table 10, the results are the same. It proves that we can use volume as a control variable for time-series regression to capture cross-sectional effects in both datasets and different spread estimators.

4.2. Coin analysis

In this section, we analyze nine coins separately. We have several hypotheses about what can affect spread in cryptocurrency markets:

1. Public interest to a particular currency. Changes in interest for cryptocurrencies affect demand and supply curves. Therefore, it may affect the spread. We use google trend values as a proxy for it.
2. Upwards or downwards changes in prices. Significant movements in price usually increase the spread. We check the absolute values of returns.
3. Trading volume. As we already mentioned and proved, the volume correlates with the spread.
4. Supply of coins. Bitcoin and other minable currencies have the expected number of coins produced each day, but fluctuations in computational power may lead to unexpected variations in supply and therefore affect spread.
5. Bitcoin spread. Bitcoin liquidity may affect other currencies' liquidity.
6. Gold spread. Media often call Bitcoin as digital gold (Zigah 2020), so it is worth to check the connection between bitcoin and gold. Previous research claims there is no connection in returns (Liu et al. 2018), so we do not expect to find a connection in liquidity as well.
7. VIX value. VIX tends to increase during emerging market crisis periods so that it can be used as a proxy for global liquidity for risky assets (Matsumoto et al. 2011).

To run a regression on time-series, we have to make sure all variables are stationary. First, we find an order of integration for spread time-series of every currency. To do it, we run the Augmented Dickey-Fuller test with null of non-stationarity and KPSS test with null of stationarity. We reject the null hypothesis at 5% significance level.

Second, we check if our factors' time-series are stationary on the same integration level as spreads. We exclude currencies where we could not make both factors and spread stationary. Then we run the following regression for every currency:

$$Spread_t = \alpha_t + \beta * Factor_t + \varepsilon_t$$

It is crucial to notice that betas may be upward-biased because of omitted variable bias. To solve this issue as the last step, we run the regression for the model, including all significant factors from previous steps. We also perform additional checks for multicollinearity for this model by calculating VIF value for every factor. We use Generalized Least Squares regression to cope with possible autocorrelation in errors (Aitken 1936).

We also run regression on lagged factors to check if we can use these factors for the prediction of future values of spread. If this regression shows good results, we run Vector Autoregression with one lag and perform Granger-Causality tests to understand relationships better.

In the following Table 11, we provide an integration order for spread based on the Augmented Dickey-Fuller test and KPSS test.

Table 11. Integration order of spread for different currencies

	p-value (ADF)	Integration order	p-value (KPSS)	Integration order
Bitcoin	0.0023	I(1)	0.1341	I(0)
Dash	0.0319	I(0)	0.2457	I(0)
EOS	0.0001	I(0)	0.3249	I(0)
Ethereum	0.0149 or 0.0016	I(4) or I(2) for log	0.0598	I(0)
Litecoin	0.0321	I(0)	0.7752	I(0)
Monero	0.0163	I(0)	0.4320	I(1)
Ripple	0.0024	I(0)	0.3358	I(0)
Zcash	0.0122	I(0)	0.1908	I(0)
Ethereum Classic	0.0006	I(1)	0.4905	I(0)

Results are contradicting because of different null hypotheses and small sample sizes, so for further research, we assume integration order of one for spread, as it includes all currencies for both tests.

4.2.1. Public interest

As a proxy for a public interest, we use Google Trend Value. We check the first difference and find all series are stationary except Monero. Therefore, we run eight regressions on google trend value for every currency except Monero (Table 12). It shows relatively good results for time-series with an adjusted R^2 of about 30%. EOS and Zcash are exceptions: they do not have a relation with Google Trend and have low values of R^2.

Table 12. GLS regression of Spread on Google Trend.

	const	Δ Google Trend$_t$	Adjusted R^2
Bitcoin	-5.394e-05 (7.62e-05)	0.0093*** (0.002)	0.336
Dash	-0.0005 (0.000)	0.0377* (0.018)	0.189
EOS	-1.39e-05 (4.54e-05)	-8.188e-07 (0.000)	-0.053
Ethereum	1.571e-07 (0.000)	0.0134*** (0.003)	0.297
Litecoin	-0.0001 (0.000)	0.0242*** (0.002)	0.515
Ripple	-6.474e-05 (0.000)	0.0208*** (0.002)	0.351
Zcash	4.686e-05 (0.000)	0.0037 (0.003)	0.028
Ethereum Classic	-0.0002 (0.000)	0.0106* (0.004)	0.362

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 13. GLS regression of spread on lagged Google Trend.

	const	Δ Google Trend$_{t-1}$	Adjusted R^2
Bitcoin	-4.445e-06 (3.42e-05)	0.0058** (0.002)	0.675
Dash	-0.0003 (0.000)	0.0278*** (0.007)	0.521
EOS	-2.332e-06 (4.06e-05)	0.0005 (0.001)	-0.032
Ethereum	1.711e-05 (0.000)	0.0045** (0.002)	0.404
Litecoin	6.627e-05 (8.33e-05)	0.0223*** (0.001)	0.906
Ripple	-9.932e-05 (0.000)	0.0039*** 0.001	0.423
Zcash	1.925e-05 (0.000)	-0.0014 (0.002)	-0.040

Ethereum Classic	-7.019e-05 (0.000)	0.0105* (0.004)	0.488

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Regression on lagged google trend shows the same patterns with insignificant coefficients for EOS and Zcash, but higher R^2 values (Table 13). To check the direction of the relationship between spread and google trend, we perform an f-test on granger-causality.

Table 14. Granger-causality test between Google Trend and Spread

	Google Trend Granger-cause Spread		Spread Granger-cause Google Trend	
	p-value	Reject null	p-value	Reject null
Bitcoin	0.002	Yes	0.385	No
Dash	0.628	No	0.713	No
EOS	0.762	No	0.008	Yes
Ethereum	0.170	No	0.694	No
Litecoin	0.356	No	0.754	No
Ripple	0.569	No	0.814	No
Zcash	0.931	No	0.313	No
Ethereum Classic	0.673	No	0.655	No

We can separate currencies in two groups: currencies with no relation to google trend or non-stationary google trend and others. The first group includes EOS, Monero, and Zcash. The second group includes Bitcoin, Dash, Ethereum, Litecoin, Ripple, Zcash, and Ethereum Classic.

The first group shows no relation to google trend and its lagged version with R^2 close to zero. At the same time, the second group has significant coefficients near the google trend and lagged google trend. Their average R^2 is equal to around 30% and 50%, respectively.

4.2.2. Changes in prices

To measure changes in prices, we use the absolute value of returns. Based on stationarity tests, we conclude that the first difference of the absolute value of returns is stationary for every currency except Zcash and Litecoin. Therefore, we run regressions for seven currencies. Changes in prices have almost no effect on the spread (Table 15 and Table 16). Only Monero coin shows some connection, but with very low R^2.

Table 15. GLS regression of spread on absolute returns.

	const	Δ $AbsReturn_t$	Adjusted R^2
Bitcoin	-0.0002 (0.000)	-0.0021 (0.012)	-0.044
Dash	-0.0007 (0.001)	-0.0154 (0.031)	-0.034
EOS	-1.302e-05 (4.22e-05)	0.0013 (0.004)	-0.043
Ethereum	-0.0002 (0.000)	-0.0118 (0.018)	-0.019
Monero	2.518e-05 (8.99e-05)	0.0174* (0.007)	0.133
Ripple	-0.0003 (0.000)	-0.0236 (0.019)	0.046
Ethereum Classic	-0.0005 (0.000)	0.0143 (0.021)	-0.029

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 16. GLS regression of spread on lagged absolute returns.

	const	Δ $AbsReturn_{t-1}$	Adjusted R^2
Bitcoin	-0.0002 (0.000)	-0.0113 (0.013)	-0.014
Dash	-0.0007 (0.001)	-0.0247 (0.050)	-0.024
EOS	-2.332e-06 (4.06e-05)	0.0037 (0.003)	0.051
Ethereum	-0.0002 (0.000)	-0.0165 (0.029)	0.010
Monero	1.365e-05 (8.77e-05)	-0.0191** (0.006)	0.134
Ripple	-0.0003 (0.000)	-0.0152 (0.029)	-0.005
Ethereum Classic	-0.0004 (0.0000	-0.0362 (0.033)	0.051

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.3. Volume

We already know that volume can explain cross-sectional differences among cryptocurrencies, but it also makes sense to check time-series relation. The first difference of logarithm of the volume is stationary for all currencies except Litecoin. Therefore, we run eight regressions. Different coins show different picture regarding connection with Volume (Table 17). It may happen because of measurement errors in volume, so we can try an instrumental variable approach to cope with this issue. Lagged volume has no significance for spread (Table 18).

Table 17. GLS regression of spread on the logarithm of volume.

	const	Δ Ln($Volume_t$)	Adjusted R^2
Bitcoin	-5.731e-05 (7.32e-05)	0.0006* (0.000)	0.292
Dash	-0.0006 (0.000)	0.0010 (0.001)	0.078
EOS	9.317e-06 (4.24e-05)	0.0001 (8.1e-05)	0.127
Ethereum	-4.591e-05 (0.000)	0.0007** (0.000)	0.240
Monero	2.591e-05 (9.82e-05)	0.0001 (0.000)	-0.020
Ripple	-0.0002 (0.000)	0.0006* (0.000)	0.152
Zcash	7.178e-05 (9.78e-05)	0.0003** (0.000)	0.199
Ethereum Classic	-0.0003 (0.000)	0.0010 (0.001)	0.159

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 18. GLS regression of spread on the lagged logarithm of volume.

	const	Δ $BTCSpread_t$	Adjusted R^2
Dash	-0.0007 (0.001)	-0.5313 (1.279)	-0.007
EOS	-1.225e-05 (4.06e-05)	-0.8676 (1.032)	-0.011
Ethereum	-0.0002 (0.000)	-0.0780 (0.429)	-0.039
Litecoin	-0.0004 (0.000)	-0.4092 (0.616)	0.050
Monero	8.979e-06 (9.54e-05)	1.6509 (1.802)	-0.025
Ripple	-0.0003 (0.000)	-0.1480 (0.440)	-0.020
Zcash	3.668e-05 (9.62e-05)	-5.9538** (1.737)	0.220
Ethereum Classic	-0.0005 (0.000)	-0.5028 (0.793)	0.041

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.4. Coins Supply

Coins have some expected supply, but due to the variations in computer power, realized supply might differ from expected. In coins managed by private companies, these variations may happen because of unexpected decision to issue more coins. The first difference of supply for all coins except Ripple is stationary, so we run eight regressions.

We can see that supply does not affect liquidity (Table 19 and Table 20). It happens because circulating supply is so large that small fluctuations in coins production can show any significant influence on liquidity.

Table 19. GLS regression of spread on the supply of coins.

	const	Δ Supply$_t$	Adjusted R^2
Bitcoin	-0.0002 (0.000)	-2.062e-07 (7.61e-07)	-0.045
Dash	-0.0007 (0.001)	1.946e-06 (1.84e-06)	-0.022
EOS	7.519e-06 (3.84e-05)	-1.991e-12 (5.3e-11)	-0.055
Ethereum	-0.0002 (0.000)	-2.298e-08 (8.48e-08)	-0.045
Litecoin	-0.0004 (0.000)	-1.059e-07 (5.63e-08)	-0.033
Monero	2.684e-05 (0.000)	1.608e-07 (1.65e-06)	-0.052
Zcash	-1.567e-05 (0.000)	1.205e-07 (8.46e-08)	0.023
Ethereum Classic	8.68e-09 (3.88e-08)	-0.0004 (0.000)	-0.045

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 20. GLS regression of spread on the lagged supply of coins.

	const	Δ Supply$_{t-1}$	Adjusted R^2
Bitcoin	-0.0001 (0.000)	1.978e-06 (1.54e-06)	0.037
Dash	-0.0007 (0.001)	5.259e-07 (3.74e-06)	-0.044
EOS	8.576e-06 (3.73e-05)	9.818e-11 (2.29e-11)	0.033
Ethereum	1.555e-07 (2.09e-07)	-0.0001 (0.000)	-0.017
Litecoin	-0.0004 (0.000)	-1.323e-07 (7.03e-08)	-0.026
Monero	-3.823e-05 (0.000)	-5.818e-07 (1.5e-06)	-0.045
Zcash	-1.597e-05 (0.000)	1.213e-07 (8.82e-08)	0.024
Ethereum Classic	-0.0005	-7.427e-09	-0.045

	(0.000)	(1.62e-08)	

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.5. Bitcoin Spread

As we already checked (Table 11), the first difference of spread is stationary for all currencies, but we have to exclude bitcoin. Therefore, we run eight regressions of spread on bitcoin spread. We can see that most of the variations in the spread can be explained by variations in Bitcoin spread (Table 21). It means cryptocurrencies share many common liquidity factors with bitcoin. However, EOS, Monero, and Zcash are among exceptions again. They show no or low connection with bitcoin. Lagged bitcoin spread is also insignificant for today's spread (Table 22). Only Zcash's coefficient is significant.

Table 21. GLS regression of Spread on BTC Spread.

	const	Δ $BTCSpread_t$	Adjusted R^2
Dash	-2.058e-05 (0.000)	4.6329*** (0.110)	0.955
EOS	-1.56e-05 (3.62e-05)	2.2287** (0.640)	0.210
Ethereum	8.029e-05 (8.02e-05)	1.5603*** (0.116)	0.840
Litecoin	-2.03e-05 (5.1e-05)	2.2528*** (0.221)	0.946
Monero	1.489e-05 (9.03e-05)	-3.6446 (1.975)	0.077
Ripple	-7.882e-06 (3.24e-05)	1.5953*** (0.034)	0.972
Zcash	2.066e-05 (9.49e-05)	6.2489* (2.217)	0.234
Ethereum Classic	-5.627e-06 (8.47e-05)	2.8936*** (0.261)	0.928

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 22. GLS regression of spread on lagged BTC Spread.

	const	Δ $BTCSpread_t$	Adjusted R^2
Dash	-0.0007 (0.001)	-0.5313 (1.279)	-0.007
EOS	-1.225e-05 (4.06e-05)	-0.8676 (1.032)	-0.011
Ethereum	-0.0002 (0.000)	-0.0780 (0.429)	-0.039
Litecoin	-0.0004 (0.000)	-0.4092 (0.616)	0.050
Monero	8.979e-06	1.6509	-0.025

	(9.54e-05)	(1.802)	
Ripple	-0.0003 (0.000)	-0.1480 (0.440)	-0.020
Zcash	3.668e-05 (9.62e-05)	-5.9538** (1.737)	0.220
Ethereum Classic	-0.0005 (0.000)	-0.5028 (0.793)	0.041

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.6. Gold Spread

The first difference of gold spread is stationary, so we can run nine regressions to check the existence or absence of relation between gold and cryptocurrencies. We also check the lagged values of gold spread. Empirical results confirm the hypothesis that cryptocurrencies, as it was expected, have no exposure to gold (Table 23 and Table 24).

Table 23. GLS regression of spread on the logarithm of gold spread.

	const	Δ Ln(GoldSpread$_t$)	Adjusted R^2
Bitcoin	-9.393e-05 (9.84e-05)	-0.1255 (0.117)	-0.042
Dash	-0.0004 (0.000)	-0.9331 (0.753)	-0.033
EOS	-2.055e-05 (3.78e-05)	-0.2762 (0.159)	0.131
Ethereum	-7.942e-05 (0.000)	-0.1365 (0.157)	-0.046
Litecoin	-0.0002 (0.000)	-0.3837 (0.303)	-0.035
Monero	0.1729 (0.367)	1.626e-05 (9.46e-050	-0.039
Ripple	-0.0002 (0.000)	-0.2097 (0.209)	-0.042
Zcash	2.179e-05 (0.000)	-0.1520 (0.383)	-0.045
Ethereum Classic	-0.0002 (0.000)	-0.8061 (0.426)	-0.017

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 24. GLS regression of spread on the lagged logarithm of gold spread.

	const	Δ GoldSpread$_{t-1}$	Adjusted R^2
Dash	-8.87e-05 (9.47e-05)	0.0659 (0.050	-0.038
EOS	-0.0004 (0.000)	0.4238 (0.359)	-0.030
Ethereum	-7.481e-05	0.0489	-0.046

	(0.000)	(0.089)	
Litecoin	-0.0002 (0.000)	0.1325 (0.167)	-0.039
Monero	6.964e-07 (8.75e-05)	-0.2702 (0.093)	0.140
Ripple	-0.0002 (0.000)	0.0385 (0.094)	-0.046
Zcash	2.861e-05 (0.000)	0.0747 (0.199)	-0.042
Ethereum Classic	-0.0002 (0.000)	0.1302 (0.265)	-0.043

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.7. VIX

VIX index is a significant macroeconomic indicator. It represents the expected volatility of the market and can be used as a proxy for overall liquidity for risky assets in the market. Its first difference is stationary, so that we can run nine regressions. Empirical results confirm that cryptocurrencies' liquidity has no connection with macroeconomic factors, as it was expected (Table 25 and Table 26).

Table 25. GLS regression of Spread on VIX index.

	const	Δ VIX_t	Adjusted R^2
Bitcoin	-9.201e-05 (9.72e-05)	9.124e-06 (9.39e-06)	-0.043
Dash	-0.0004 (0.000)	5.083e-05 (4.9e-05)	-0.041
EOS	-1.623e-05 (4.16e-05)	9.063e-06 (1.06e-05)	-0.025
Ethereum	-7.832e-05 (0.000)	5.299e-05 (4.26e-05)	-0.002
Litecoin	-0.0002 (0.000)	1.081e-05 (2.2e-05)	-0.046
Monero	2.726e-05 (8.86e-05)	-5.854e-05* (2.64e-05)	0.160
Ripple	-0.0002 (0.000)	1.515e-05 (1.47e-05)	-0.043
Zcash	2.588e-05 (0.000)	-1.596e-06 (3.77e-05)	-0.053
Ethereum Classic	-0.0002 (0.000)	1.905e-05 (2.98e-05)	-0.045

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

Table 26. GLS regression of spread on lagged VIX index.

	const	Δ VIX$_{t-1}$	Adjusted R^2
Bitcoin	-8.106e-05 (8.57e-05)	2.928e-05 (2.83e-05)	-0.006
Dash	-0.0004 (0.000)	0.0001 (0.000)	-0.003
EOS	-1.117e-05 (4.12e-05)	1.215e-05 (1.72e-05)	-0.007
Ethereum	-5.783e-05 (0.000)	5.258e-05 (4.74e-05)	-0.008
Litecoin	-0.0002 (0.000)	6.324e-05 (5.99e-05)	-0.003
Monero	2.32e-05 (8.97e-05)	4.983e-05* (1.97e-05)	0.090
Ripple	-0.0001 (0.000)	4.522e-05 (4.67e-05)	-0.011
Zcash	2.729e-05 (0.000)	8.171e-06 (4.7e-05)	-0.050
Ethereum Classic	-0.0002 (0.000)	0.0001 (7.79e-05)	0.017

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

4.2.8. Overall results

We confirm that public interest affects liquidity. Increasing the public interest to the coin leads to an increase in transaction costs, so spread increases. It happens because public interest may lead to a demand curve shift, therefore increasing spread.

We also confirm that there is a commonality on liquidity factors among cryptocurrencies, but these common factors are independent of the regular market. Results support previous works claiming the absence of cryptocurrencies' exposure to metals and macro factors.

Another unexpected result is the unique case of Monero, Zcash, and EOS. They have almost no connection with other markets. It means they have other sources of liquidity. One possible explanation of this phenomenon is the support of anonymous transactions by Monero and Zcash. An empirical study shows that Zcash has more exposure to the other currencies' liquidity than Monero. We can see it as a supporting argument because Zcash allows both anonymous and non-anonymous transactions. Therefore, we can conclude that Zcash has a connection with market liquidity through a channel of non-anonymous transactions. EOS case is more difficult because it positioned not as a currency but as a platform for developing apps on the blockchain.

4.3. Time-series panel analysis

To confirm the results in the previous section, we analyze the group of cryptocurrencies together. We exclude Monero, Zcash, and EOS as they showed different behavior relative to other currencies, so we have five currencies in our panel data. As we use first differences to ensure stationarity of variables, we run pooled OLS regressions.

Table 27. First difference regression currencies' CRSP spread on bitcoin spread and google trend.

	Model 1	Model 2	Model 3
BTCSpread	-	2.4949*** (0.4887)	2.4047*** (0.4349)
GoogleTrend	0.1810** (0.0653)	-	0.0455* (0.0202)
const	-0.0001*** (3.729e-05)	8.793e-06 (1.551e-05)	1.822e-05 (1.315e-05)
R^2	0.1428	0.7704	0.7784
R^2 Between	-0.2614	2.22e-16	-0.0424
R^2 Within	0.1464	0.7772	0.7857

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

To verify our results, we also run the same regression on the bigger dataset from BitFinex Api (Table 28). Results are consistent with our previous findings.

Table 28. First difference regression currencies' CHL spread on bitcoin spread and google trend.

	Model 1	Model 2	Model 3
BTCSpread	-	0.8343*** (0.0832)	0.8121*** (0.0854)
GoogleTrend	0.2964*** (0.0859)	-	0.0793* (0.0368)
const	-0.0003 (0.0001)	7.357e-06 (8.316e-05)	8.046e-05 (9.079e-05)
R^2	0.0303	0.2292	0.2312
R^2 Between	-0.9021	1.11e-16	0.0229
R^2 Within	0.0317	0.2296	0.2315

P-values: *** - < 0.001, ** - < 0.01, * - < 0.05

5. Portfolio optimization

After having analyzed different estimators of liquidity and factors affecting liquidity, we moved on to constructing portfolios. Our principal aim in the portfolio construction was to test how transaction costs may affect portfolio performance because transaction costs mean rebalance costs.

There are to parts in the portfolio optimization problem. The first part is about selecting the best distribution of the assets in the beginning. It means setting target weights for every asset in the portfolio. The second is about maintaining this distribution over time. We can achieve it by rebalancing the portfolio by specific rules.

We used the same period of two years as for the previous analysis. We use monthly returns and spreads for nine currencies. We construct three portfolios: equally weighted, size weighted, and liquidity weighted. For a liquidity-weighted portfolio, we define liquidity as the inverse of spread because big spread means low liquidity. We include a minimum of one percent of every currency. In the following table, we provide weights for these portfolios (Table 29). In both size-weighted and liquidity-weighted most of the weight goes to Bitcoin and Ethereum. However, in a size-weighted portfolio, Bitcoin has maximum weight while in a liquidity-weighted portfolio, Ethereum has maximum weight.

Table 29. Portfolio weights

	Equally-weighted	Size-weighted	Liquidity-weighted
Bitcoin	12 %	57 %	39 %
Dash	11 %	1 %	1 %
EOS	11 %	3 %	4 %
Ethereum	11 %	23 %	48 %
Litecoin	11 %	2 %	3 %
Monero	11 %	1 %	1 %
Ripple	11 %	11 %	2 %
Zcash	11 %	1 %	1 %
Ethereum Classic	11 %	1 %	1 %

For the second part of the problem, we define half of the spread as the transaction costs and ignore any other costs like commissions. For each of the portfolios, we use several approaches to rebalance. In the case of no rebalance, we have no exposure to the transaction costs, but the portfolio may have significant deviations from our target weights. Another option is to rebalance the portfolio every fixed period. We use periods of one month, three months, and half a year. For every portfolio, we calculate the Sharpe ratio and compare them. We assume risk-free rate is 0 %.

To rebalance our portfolio, we need to solve the system of linear equations for a_i:

$$w_i = \frac{a_i}{\sum b_j - \sum sign_j * (a_j - b_j) * c_j} \ for\ every\ currency\ i, where$$

a_i $- value\ of\ the\ currecy\ i\ after\ rebalance$

b_i $- value\ \ of\ the\ currency\ i\ before\ rebalance$

$w_i - target\ weight$

$c_i - transaction\ costs$

$sign_i - 1\ if\ current\ weight\ is\ less\ than\ target\ weight\ and\ (-1)\ otherwise$

Therefore, we have asset value after rebalancing in the numerator and portfolio value after rebalancing in the divisor. We can transform these equations and then solve the system of linear equations:

$$w_i * \left(\sum b_j + \sum sign_j * b_j * c_j\right) = w_i * \left(\sum sign_j * a_j * c_j\right) + a_i$$

Table 30. Sharpe ratios for different portfolios and different periods for rebalancing

	No rebalance	One month	Three months	Half-year
Equally-weighted	-0.0265	0.0111	0.0058	0.0089
Size-weighted	0.0757	0.0794	0.078	0.0795
Liquidity-weighted	0.0429	0.0456	0.0481	0.0443

There are two important conclusions from these results. First, we can see that the size-weighted portfolio shows better performance than the liquidity-weighted portfolio. It is one more evidence that investor decision in the cryptocurrency market does not depend on liquidity. Second, to rebalance the portfolio is always better than just keep, but due to the transaction costs, rarer rebalancing can give us better results sometimes.

6. Conclusion

To conclude it all, liquidity plays a vital role in any market, including the cryptocurrency market. However, the topic of liquidity in the cryptocurrency market is not well covered due to the youthfulness crypto industry. In this paper, we explore the possibility of using different liquidity estimators in the cryptocurrency market and examine factors that may affect liquidity.

From our analysis of liquidity estimators, we conclude that we cannot use typical estimators for stock markets. Empirical analysis shows a low correlation of low-frequency estimators based on price with benchmark estimator based on quoted bid and ask. These estimators have several assumptions regarding prices, which do not hold in the crypto world. For example, CHL and HL estimators assume Geometric Brownian Motion for prices, which is not valid for cryptocurrencies. They have strong negative autocorrelation in returns due to the boom-bust speculative cycles, and underlying value is harder to determine for investors.

From Fama-Macbeth cross-sectional analysis, we find factors that influence on liquidity variations between different cryptocurrencies. Different levels of public interest can explain some differences in the liquidity of currency, but not all. The best proxy for variations between different currencies is trading volume. It incorporates public interest and some other unknown factors.

We test seven hypotheses regarding factors that may affect cryptocurrency liquidity in the time-series dimension. The empirical analysis confirms the absence of cryptocurrencies' exposure to metals and macro factors. We find a high level of commonality in liquidity among cryptocurrencies by checking the relation between bitcoin's spread and other currencies' spreads. However, some currencies have no or very low connection with the rest of the market. One possible hypothesis explaining it is the ability of these currencies to hide transaction information, and hence they have other sources of liquidity. The verification of this hypothesis may be part of the future development of this work. There is no evidence confirming the relationship between spread and coins supply, changes in prices and volume. Public interest has some effect on the spread, but this effect is almost insignificant comparing to common factors included in bitcoin spread.

The first difference panel regression confirms all the previously mentioned results. The empirical analysis of a more extensive dataset based on a more inaccurate estimator is also consistent with other results.

We investigated the influence of the period between rebalancing on the performance of different portfolios. We found out that while rebalancing shows better performance compared to the absence of rebalancing, more often rebalancing can show worse performance because of transaction costs.

There are several possible directions for the development of this work. First, developing effective spread estimators for the crypto-currency market can simplify any further research. Second, reproducing outcomes of this work on broader datasets by coinmarketcap.com will improve confidence in the results of current work. Third, closer attention to coins with different behavior like Monero can give some unexpected results and insights about the current crypto market.

To sum up, in this paper, we conducted a detailed investigation of liquidity estimators and liquidity factors for cryptocurrency markets, and we investigated the performance of different portfolios with different times between rebalancing in the presence of transaction costs.

Bibliography

Abdi, Farshid, and Angelo Ranaldo. 2017. "A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices." *Review of Financial Studies* 30 (12): 4437–80. https://doi.org/10.1093/rfs/hhx084.

Acharya, Viral V., and Lasse Heje Pedersen. 2005. "Asset Pricing with Liquidity Risk." *Journal of Financial Economics* 77 (2): 375–410. https://doi.org/10.1016/j.jfineco.2004.06.007.

Aitken, Alexander C. 1936. "IV.—On Least Squares and Linear Combination of Observations." *Proceedings of the Royal Society of Edinburgh* 55 (September): 42–48. https://doi.org/10.1017/S0370164600014346.

Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets* 5 (1): 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6.

Amihud, Yakov, and Haim Mendelson. 1986. "Asset Pricing and the Bid-Ask Spread." *Journal of Financial Economics* 17 (2): 223–49. https://doi.org/10.1016/0304-405X(86)90065-6.

Brennan, Michael, Sahn Wook Huh, and Avanidhar Subrahmanyam. 2013. "An Analysis of the Amihud Illiquidity Premium." *Review of Asset Pricing Studies* 3 (1): 133–76. https://doi.org/10.1093/rapstu/ras017.

Brockman, Paul, Dennis Y. Chung, and Christophe Prignon. 2009. "Commonality in Liquidity: A Global Perspective." *Journal of Financial and Quantitative Analysis* 44 (4): 851–82. https://doi.org/10.1017/S0022109009990123.

Buterin, Vitalik. 2014. "A Next-Generation Smart Contract and Decentralized Application Platform." *Etherum*, no. January: 1–36. https://github.com/ethereum/wiki/wiki/White-Paper.

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52 (1): 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x.

Cheng, Shuenn Ren. 2007. "A Study on the Factors Affecting Stock Liquidity." *International Journal of Services and Standards* 3 (4): 453–75. https://doi.org/10.1504/IJSS.2007.015227.

Choi, J. Y., Dan Salandro, and Kuldeep Shastri. 1988. "On the Estimation of Bid-Ask Spreads: Theory and Evidence." *The Journal of Financial and Quantitative Analysis* 23 (2): 219. https://doi.org/10.2307/2330882.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2000. "Commonality in Liquidity." *Journal of Financial Economics* 56 (1): 3–28. https://doi.org/10.1016/S0304-405X(99)00057-4.

Chung, Kee H., and Hao Zhang. 2014. "A Simple Approximation of Intraday Spreads Using Daily Data." *Journal of Financial Markets* 17 (1): 94–120. https://doi.org/10.1016/j.finmar.2013.02.004.

Copeland, Thomas E., and Dan Galai. 1983. "Information Effects on the Bid-Ask Spread." *The Journal of Finance* 38 (5): 1457–69. https://doi.org/10.1111/j.1540-6261.1983.tb03834.x.

Corwin, Shane A., and Paul Schultz. 2012. "A Simple Way to Estimate Bid-Ask Spreads from Daily High and Low Prices." *Journal of Finance* 67 (2): 719–60. https://doi.org/10.1111/j.1540-6261.2012.01729.x.

Demsetz, Harold. 1968. "The Cost of Transacting." *The Quarterly Journal of Economics* 82 (1): 33. https://doi.org/10.2307/1882244.

Eleswarapu, Venkat R. 1997. "Cost of Transacting and Expected Returns in the Nasdaq Market." *Journal of Finance* 52 (5): 2113–27. https://doi.org/10.1111/j.1540-6261.1997.tb02754.x.

"FAQ about Google Trends Data - Trends Help." n.d. Accessed May 6, 2020. https://support.google.com/trends/answer/4365533.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka. 2009. "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics* 92 (2): 153–81. https://doi.org/10.1016/j.jfineco.2008.06.002.

Harris, Larry, and Andrea Amato. 2019. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects: A Replication." *Critical Finance Review* 8 (1–2): 173–202. https://doi.org/10.1561/104.00000058_supp.

Hasbrouck, Joel. 2004. "Liquidity in the Futures Pits: Inferring Market Dynamics from Incomplete Data." *Journal of Financial and Quantitative Analysis* 39 (2): 305–26. https://doi.org/10.1017/s0022109000003082.

Hasbrouck, Joel, and Duane J. Seppi. 2001. "Common Factors in Prices, Order Flows, and Liquidity." *Journal of Financial Economics* 59 (3): 383–411. https://doi.org/10.1016/s0304-405x(00)00091-x.

Holden, Craig W. 2009. "New Low-Frequency Spread Measures." *Journal of Financial Markets* 12 (4): 778–813. https://doi.org/10.1016/j.finmar.2009.07.003.

Huberman, Gur, and Dominika Halka. 2001. "Systematic Liquidity." *Journal of Financial Research* 24 (2): 161–78. https://doi.org/10.1111/j.1475-6803.2001.tb00763.x.

Kappos, George, Haaroon Yousaf, Mary Maller, and Sarah Meiklejohn. 2018. "An Empirical Analysis of Anonymity in Zcash." In *27th USENIX Security Symposium (USENIX Security 18)*, 463–77. Baltimore, MD: {USENIX} Association. https://www.usenix.org/conference/usenixsecurity18/presentation/kappos.

Lesmond, David A, Joseph P Ogden, and Charles A Trzcinka. 1999. "A New Estimate of Transaction Costs." *Review of Financial Studies* 12 (5): 1113–41. https://doi.org/10.1093/rfs/12.5.1113.

Liu, Yukun, Aleh Tsyvinski, Andrew Atkeson, Nicola Borri, Eduardo Davila, Stefano Giglio, William Goetzmann, Stephen Roach, and Robert Shiller. 2018. "Risks and Returns of Cryptocurrency." NBER Working Papers. http://www.nber.org/papers/w24877.

Liu, Yukun, Aleh Tsyvinski, Xi Wu, We Thank, Nicola Borri, Markus Brunnermeier, Kent Daniel, et al. 2019. "Common Risk Factors in Cryptocurrency." NBER Working Paper. http://www.nber.org/papers/w25882.

Matsumoto, Akito, Menzie Chinn, Stijn Claessens, Julian Di Giovanni, Christian Echeverria, Anna Ivanova, Bin Li, et al. 2011. "Global Liquidity: Availability of Funds for Safe and Risky Assets IMF Working Paper Research Department Global Liquidity: Availability of Funds for Safe and Risky Assets 1."

Nakamoto, Satoshi. 2008. "Bitcoin: A Peer-to-Peer Electronic Cash System." 2008. https://bitcoin.org/bitcoin.pdf.

Pástor, Ľuboš, and Robert F Stambaugh. 2003. "Liquidity Risk and Expected Stock Returns."

Journal of Political Economy 111 (3): 642–85. https://doi.org/10.1086/374184.

Roll, Richard. 1984. “A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market.” *The Journal of Finance* 39 (4): 1127–39. https://doi.org/10.1111/j.1540-6261.1984.tb03897.x.

Shen, D., Andrew Urquhart, and Pengfei Wang. 2019. “A Three-Factor Pricing Model for Cryptocurrencies.” *Finance Research Letters*. https://doi.org/10.1016/j.frl.2019.07.021.

Urquhart, Andrew. 2016. “The Inefficiency of Bitcoin.” *Economics Letters* 148 (November): 80–82. https://doi.org/10.1016/j.econlet.2016.09.019.

Wang, Sha, and Jean Philippe Vergne. 2017. “Buzz Factor or Innovation Potential: What Explains Cryptocurrencies’ Returns?” *PLoS ONE* 12 (1). https://doi.org/10.1371/journal.pone.0169556.

Wei, Wang Chun. 2018. “Liquidity and Market Efficiency in Cryptocurrencies.” *Economics Letters* 168 (July): 21–24. https://doi.org/10.1016/j.econlet.2018.04.003.

Zigah, Xolali. 2020. “Council Post: Is Bitcoin Really Digital Gold?” *Forbes*, 2020. https://www.forbes.com/sites/forbesfinancecouncil/2020/05/11/is-bitcoin-really-digital-gold/.

Summary

Nowadays, in the modern financial markets with a wide variety of different investment opportunities, liquidity became a vital decision indicator. The impact of liquidity on the choice of investment draws more and more attention. Liquidity is often defined as "an ability to trade large quantities quickly at low cost with little price impact."

Cryptocurrencies have less history than any other traditional asset, but they draw significant attention from investors and researches. Since the introduction of bitcoin in 2009, they have become increasingly popular. Cryptocurrency is a digital token that can be exchanged online. It uses cryptographic hashing and digital signatures to verify transactions and avoid double-spending of the same token. These features prevent from creating tokens from the air, so the supply is limited. In this aspect, cryptocurrencies are like metals or other commodities. Because the limited quantity of cryptocurrencies is protected by the cryptography embedded in their open-source code, cryptocurrencies can potentially become valuable.

There is a wide range of papers written on the topic of liquidity. They investigate liquidity measures, construct liquidity factors, and measure their impact on asset prices. Cryptocurrencies also raise interest, so despite the relative youthfulness, there are many papers on this topic. However, there are almost no investigation on the topic of liquidity in the crypto market and its connection with different factors, including traditional markets. Cryptocurrencies have some peculiar properties that can prevent us from using traditional techniques for liquidity estimating.

The main aim of this paper is to fill this gap. Therefore, the objectives of this research are the following:

- Determine the range of existing liquidity proxies in the absence of quote data
- Validate proxies
- Explore factors affecting the liquidity of the cryptocurrency market
- Investigate how liquidity affects portfolio optimization and rebalancing strategy

This investigation stands out among others, due to the following reasons. First, we explore and compare different liquidity estimators based on low-frequency data. It allows us to use more available data sources without restrictions. Second, we check the performance of the estimators in the cryptocurrency markets and point out the limitations of existing estimators for the crypto market. Third, we run a comprehensive investigation of different factors affecting cryptocurrencies' liquidity. Finally, we explore how transaction costs may affect different portfolios' performance.

We collect daily OHLCV data for the period from 01.01.2018 to 01.03.2020. We use the BitFinex cryptocurrency exchange with provided open API and Bloomberg as data sources. From BitFinex API, we select 20 cryptocurrencies with the highest trading volume, excluding stablecoins. Selected coins are Bitcoin, Ethereum, EOS, Ripple, Litecoin, Monero, Iota, Zcash, Dash, NEO, Ethereum Classic, Eidoo, OmiseGO, Bitcoin Gold, ETP, Streamr, Golem, Santiment, Qtum, and YOYOW.

Blomberg database has access only to nine currencies. They are Bitcoin, Dash, EOS, Ethereum, Ethereum Classic, Litecoin, Monero, Ripple, and Zcash. Besides OHLCV data, we also collect quoted bid, ask, and circulating supply of coins. The circulating supply of coins is the number of coins available for trading and equal to mined coins minus unreachable coins. If we do not need additional data on bid, ask, and supply from Bloomberg, and if we do not say it implicitly, we use BitFinex API data.

To include more liquidity factors, we collect daily data for spot gold prices, including OHLCV data and quoted bid and ask. For the same reasons, we collect VIX index historical daily data. We take all the mentioned data from Bloomberg.

We also use Google trend data, which represents a number of google searches for the topic. As the topic, we use currency name.

Before we start working with spreads, it is essential to verify estimators. Cryptocurrencies show different behavior from traditional assets. It means that some assumptions made by estimator's authors may not hold in cryptocurrency markets. We run our checks on Roll estimator, Corwin and Schultz HL estimator, Abdi and Ronaldo CHL estimator.

Our first check is almost equal to the check conducted by Roll to check his estimator (Roll 1984). We know that spread has to have negative cross-sectional relation with volume. Several papers point to that (Demsetz 1968; Copeland and Galai 1983). Therefore, we run the following Fama-Macbeth cross-sectional relation for every estimator.

$$Spread = c + \beta_1 Volume$$

As a result of our checks, every estimator has a significant negative relation with volume. It means these estimators have a relation to the real spread.

Our second check is measuring the relation between estimators and the benchmark of real spread. We use the CRSP spread estimator by Chung and Zhang (Chung and Zhang 2014) as the benchmark. We check both cross-sectional and time-series correlations.

We can see all estimators have an upward bias from the CRSP estimator. One possible explanation is negative autocorrelation in transaction type. We can make such a conclusion because we know about time-series momentum in cryptocurrencies (Liu et al. 2018). It is possible to suppose that the probability of getting the opposite transaction type is more than 50%. As we can see from the paper by Choi et all, it can lead to upward bias (Choi, Salandro, and Shastri 1988). Of course, it is just a hypothesis, and to check it, we need intraday tick data, which we do not have.

For cross-sectional and time-series correlation, all estimators show much weaker performance that in the paper by Abdi and Ronaldo (Abdi and Ranaldo 2017). As we already mentioned, the cryptocurrency market is inefficient and has time-series autocorrelation. However, estimators rely on the assumption of prices moving in Geometric Brownian Motion, which is not valid in the case of cryptocurrencies. It means that the use of these estimators is very limited in cryptocurrency markets.

We focus our research on the CRSP estimator because it does not depend on assumptions needed for other estimators. We also use CHL and HL estimator as they allow us to use larger data sets. Nevertheless, in the interpretation of the results, priority goes to the CRSP estimator.

First, we examine cross-sectional factors that affect the spread. We use Google Trend value, size, or market capitalization of the currency and trading volume in dollars. All variables are in logs. We test four models by running Fama-Macbeth regressions, and then we compare "R^2 between". "R^2 between" measures the goodness of fit using the time-series average of the variables. Therefore, in our case, it measures how good our variables explain cross-sectional factors.

Following the Fama-Macbeth procedure, we run T cross-sectional regressions and then take the average of the coefficients. We use four models. Model 1 includes only Google Trend, Model 2 adds Size to Model 1, Model 3 adds Volume to Model 1, and Model 4 includes all variables:

$$Ln(Spread_i) = \beta_1 + \beta_2 * Ln(GoogleTrend_i) + \beta_3 * Ln(Size_i) + \beta_4 * Ln(Volume_i)$$

Model 1 includes only Google Trend value and can explain only 37% of spread differences among currencies. In the other three models, Google Trend is insignificant. It means that Google Trend's influence is already included in Size or Volume. Therefore, we can check it by regressing size and volume on google trend. We got 43% and 38% values of R^2 between accordingly.

Model 2, 3, and 4 explain cross-sectional differences in spreads well enough. If we look at these models, we can notice multicollinearity in Model 4 because of the close connection between Size and Volume (VIF > 10 in model 4). All models have high values of R^2 between, so they can explain Spread very well. Therefore, we can use models 2 or 3, depending on the data available. It is a beneficial finding because it proves we can use volume or size as the controls in time series effects to capture entity effects.

We also verify this statement by checking other spread estimates and other datasets. We run the same regressions using CHL and HL spreads. We also run it on a bigger dataset by BitFinex. As we have no size data in this dataset, there are only two models to check. CHL and HL spreads perform worse than CRSP spread. It shows similar results, but the significance of variables may differ. These estimators are less precise; that is why we need more coins for cross-sectional regression. We use the BitFinex dataset, which includes 20 coins. The results are the same. It proves that we can use volume as a control variable for time-series regression to capture cross-sectional effects in both datasets and different spread estimators.

We have several hypotheses about what can affect spread in cryptocurrency markets:

1. Public interest to a particular currency. Changes in interest for cryptocurrencies affect demand and supply curves. Therefore, it may affect the spread. We use google trend values as a proxy for it.
2. Upwards or downwards changes in prices. Significant movements in price usually increase the spread. We check the absolute values of returns.
3. Trading volume. As we already mentioned and proved, the volume correlates with the spread.
4. Supply of coins. Bitcoin and other minable currencies have the expected number of coins produced each day, but fluctuations in computational power may lead to unexpected variations in supply and therefore affect spread.
5. Bitcoin spread. Bitcoin liquidity may affect other currencies' liquidity.
6. Gold spread. Media often call Bitcoin as digital gold (Zigah 2020), so it is worth to check the connection between bitcoin and gold. Previous research claims there is no connection in returns (Liu et al. 2018), so we do not expect to find a connection in liquidity as well.
7. VIX value. VIX tends to increase during emerging market crisis periods so that it can be used as a proxy for global liquidity for risky assets (Matsumoto et al. 2011).

To run a regression on time-series, we have to make sure all variables are stationary. First, we find an order of integration for spread time-series of every currency. To do it, we run the Augmented Dickey-Fuller test with null of non-stationarity and KPSS test with null of stationarity. We reject the null hypothesis at 5% significance level.

Second, we check if our factors' time-series are stationary on the same integration level as spreads. We exclude currencies where we could not make both factors and spread stationary. Then we run the following regression for every currency:

$$Spread_t = \alpha_t + \beta * Factor_t + \varepsilon_t$$

It is crucial to notice that betas may be upward-biased because of omitted variable bias. To solve this issue as the last step, we run the regression for the model, including all significant factors from previous steps. We also perform additional checks for multicollinearity for this model by calculating VIF value for every factor. We use Generalized Least Squares regression to cope with possible autocorrelation in errors (Aitken 1936).

We also run regression on lagged factors to check if we can use these factors for the prediction of future values of spread. If this regression shows good results, we run Vector Autoregression with one lag and perform Granger-Causality tests to understand relationships better.

We confirm that public interest affects liquidity. Increasing the public interest to the coin leads to an increase in transaction costs, so spread increases. It happens because public interest may lead to a demand curve shift, therefore increasing spread.

We also confirm that there is a commonality on liquidity factors among cryptocurrencies, but these common factors are independent of the regular market. Results support previous works claiming the absence of cryptocurrencies' exposure to metals and macro factors.

Another unexpected result is the unique case of Monero, Zcash, and EOS. They have almost no connection with other markets. It means they have other sources of liquidity. One possible explanation of this phenomenon is the support of anonymous transactions by Monero and Zcash. An empirical study shows that Zcash has more exposure to the other currencies' liquidity than Monero. We can see it as a supporting argument because Zcash allows both anonymous and non-anonymous transactions. Therefore, we can conclude that Zcash has a connection with market liquidity through a channel of non-anonymous transactions. EOS case is more difficult because it positioned not as a currency but as a platform for developing apps on the blockchain.

To confirm the results, we analyze the group of cryptocurrencies together. We exclude Monero, Zcash, and EOS as they showed different behavior relative to other currencies, so we have five currencies in our panel data. As we use first differences to ensure stationarity of variables, we run pooled OLS regressions.

To verify our results, we also run the same regression on the bigger dataset from BitFinex API. Results are consistent with our previous findings.

After having analyzed different estimators of liquidity and factors affecting liquidity, we moved on to constructing portfolios. Our principal aim in the portfolio construction was to test how transaction costs may affect portfolio performance because transaction costs mean rebalance costs.

There are to parts in the portfolio optimization problem. The first part is about selecting the best distribution of the assets in the beginning. It means setting target weights for every asset in the portfolio. The second is about maintaining this distribution over time. We can achieve it by rebalancing the portfolio by specific rules.

We used the same period of two years as for the previous analysis. We use monthly returns and spreads for nine currencies. We construct three portfolios: equally weighted, size weighted, and liquidity weighted. For a liquidity-weighted portfolio, we define liquidity as the inverse of spread because big spread means low liquidity. We include a minimum of one percent of every currency. In both size-weighted and liquidity-weighted most of the weight goes to Bitcoin and Ethereum. However, in a size-weighted portfolio, Bitcoin has maximum weight while in a liquidity-weighted portfolio, Ethereum has maximum weight.

For the second part of the problem, we define half of the spread as the transaction costs and ignore any other costs like commissions. For each of the portfolios, we use several approaches to rebalance. In the case of no rebalance, we have no exposure to the transaction costs, but the portfolio may have significant deviations from our target weights. Another option is to rebalance the portfolio every fixed period. We use periods of one month, three months, and half a year. For every portfolio, we calculate the Sharpe ratio and compare them. We assume risk-free rate is 0 %.

There are two important conclusions from these results. First, we can see that the size-weighted portfolio shows better performance than the liquidity-weighted portfolio. It is one more evidence that investor decision in the cryptocurrency market does not depend on liquidity. Second, to rebalance the portfolio is always better than just keep, but due to the transaction costs, rarer rebalancing can give us better results sometimes.

To conclude it all, liquidity plays a vital role in any market, including the cryptocurrency market. However, the topic of liquidity in the cryptocurrency market is not well covered due to the youthfulness crypto industry. In this paper, we explore the possibility of using different liquidity estimators in the cryptocurrency market and examine factors that may affect liquidity.

From our analysis of liquidity estimators, we conclude that we cannot use typical estimators for stock markets. Empirical analysis shows a low correlation of low-frequency estimators based on price with benchmark estimator based on quoted bid and ask. These estimators have several assumptions regarding prices, which do not hold in the crypto world. For example, CHL and HL estimators assume Geometric Brownian Motion for prices, which is not valid for cryptocurrencies. They have strong negative autocorrelation in returns due to the boom-bust speculative cycles, and underlying value is harder to determine for investors.

From Fama-Macbeth cross-sectional analysis, we find factors that influence on liquidity variations between different cryptocurrencies. Different levels of public interest can explain some differences in the liquidity of currency, but not all. The best proxy for variations between different currencies is trading volume. It incorporates public interest and some other unknown factors.

We test seven hypotheses regarding factors that may affect cryptocurrency liquidity in the time-series dimension. The empirical analysis confirms the absence of cryptocurrencies' exposure to metals and macro factors. We find a high level of commonality in liquidity among cryptocurrencies by checking the relation between bitcoin's spread and other currencies' spreads. However, some currencies have no or very low connection with the rest of the market. One possible hypothesis explaining it is the ability of these currencies to hide transaction information, and hence they have other sources of liquidity. The verification of this hypothesis may be part of the future development of this work. There is no evidence confirming the relationship between spread and coins supply, changes in prices and volume. Public interest has some effect on the spread, but this effect is almost insignificant comparing to common factors included in bitcoin spread.

The first difference panel regression confirms all the previously mentioned results. The empirical analysis of a more extensive dataset based on a more inaccurate estimator is also consistent with other results.

We investigated the influence of the period between rebalancing on the performance of different portfolios. We found out that while rebalancing shows better performance compared to the absence of rebalancing, more often rebalancing can show worse performance because of transaction costs.

There are several possible directions for the development of this work. First, developing effective spread estimators for the crypto-currency market can simplify any further research. Second, reproducing outcomes of this work on broader datasets by coinmarketcap.com will improve confidence in the results of current work. Third, closer attention to coins with different behavior like Monero can give some unexpected results and insights about the current crypto market.

To sum up, in this paper, we conducted a detailed investigation of liquidity estimators and liquidity factors for cryptocurrency markets, and we investigated the performance of different portfolios with different times between rebalancing in the presence of transaction costs.