

Introducing Coin Metrics' Trusted Volume Framework

 coinmetrics.io/introducing-coin-metrics-trusted-volume-framework

July 28, 2020

By Jon Geenty and the Coin Metrics Team

Key Takeaways

- Fake trading volumes have been a dark mark on the industry – it is difficult to find a single metrics to accurately sift through the reported numbers.
- We've taken a data driven approach to the problem and are excited to introduce a "trusted volume" metric to help identify legitimate trading volume.
- Our framework for measuring the reporting quality of an exchange is broken down into three broad categories: **volume correlation**, **web traffic analytics** and **qualitative features**.
- As of June 2020, the passing exchanges for 'trusted' spot volume include Binance (and Binance US), Bitbank, Bitfinex, bitFlyer, Bitstamp, Bittrex, CEX.IO, Coinbase, Gate.io, Gemini, itBit, Kraken, and Poloniex.

Overview

Fake trading volume is a persistent problem on crypto exchanges. With little regulatory oversight, it can be difficult to determine whether reported volume numbers are accurate or exaggerated. At Coin Metrics we've taken a data driven approach to the problem and are excited to introduce a "trusted volume" metric to help identify legitimate trading volume.

Our trusted volume metric is an aggregation of the reported volume from exchanges that we consider the most accurate and trustworthy. This is based on a combination of both quantitative and qualitative features. The analysis below is limited to spot exchanges and spot volume. This is a distinction with regards to derivative exchanges such as BitMEX and Deribit.

Our framework for measuring the reporting quality of an exchange is broken down into three broad categories: **volume correlation**, **web traffic analytics** and **qualitative features**. Each of these three categories culminates in a pass/fail test. Exchanges that pass all three measures are included in our trusted volume set of metrics.

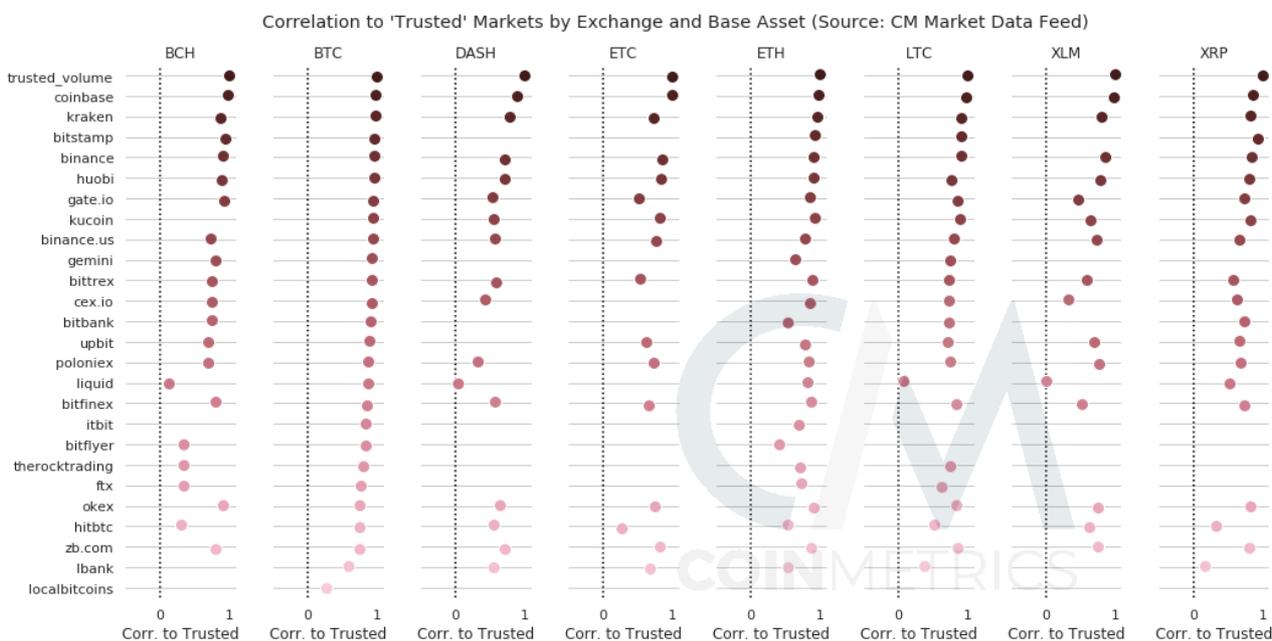
Part I: Correlation of Hourly Volume with Regulated Exchanges

During times of high volatility and large price movements, exchanges should see an increase in trading volumes in a similar magnitude to their peers. The same applies in times of low volatility and small price movement. Accordingly, authentic volume should

have a higher correlation to other authentic volume. Fake volume should surface as an outlier with lower correlation.

In this test, we took hourly volume from long-standing, regulated exchanges based in the United States and used that as our 'trusted' control group. These exchanges all have a relatively long established history of at least five years and legitimacy that has been somewhat bolstered by regulations within the United States. These exchanges also support fiat USD trading, which adds additional credibility. This list includes Bitstamp, Bittrex, Coinbase, Gemini, itBit and Kraken. All data in this section is from trading volume between June 1 and June 30, 2020.

Using a time series of the hourly volume from all markets on all exchanges in our test set, we aggregated the volume by base asset and calculated the correlation between the exchange and the volume for the same base asset on the trusted exchanges. We did this for all assets listed on *at least two or more* trusted exchanges.

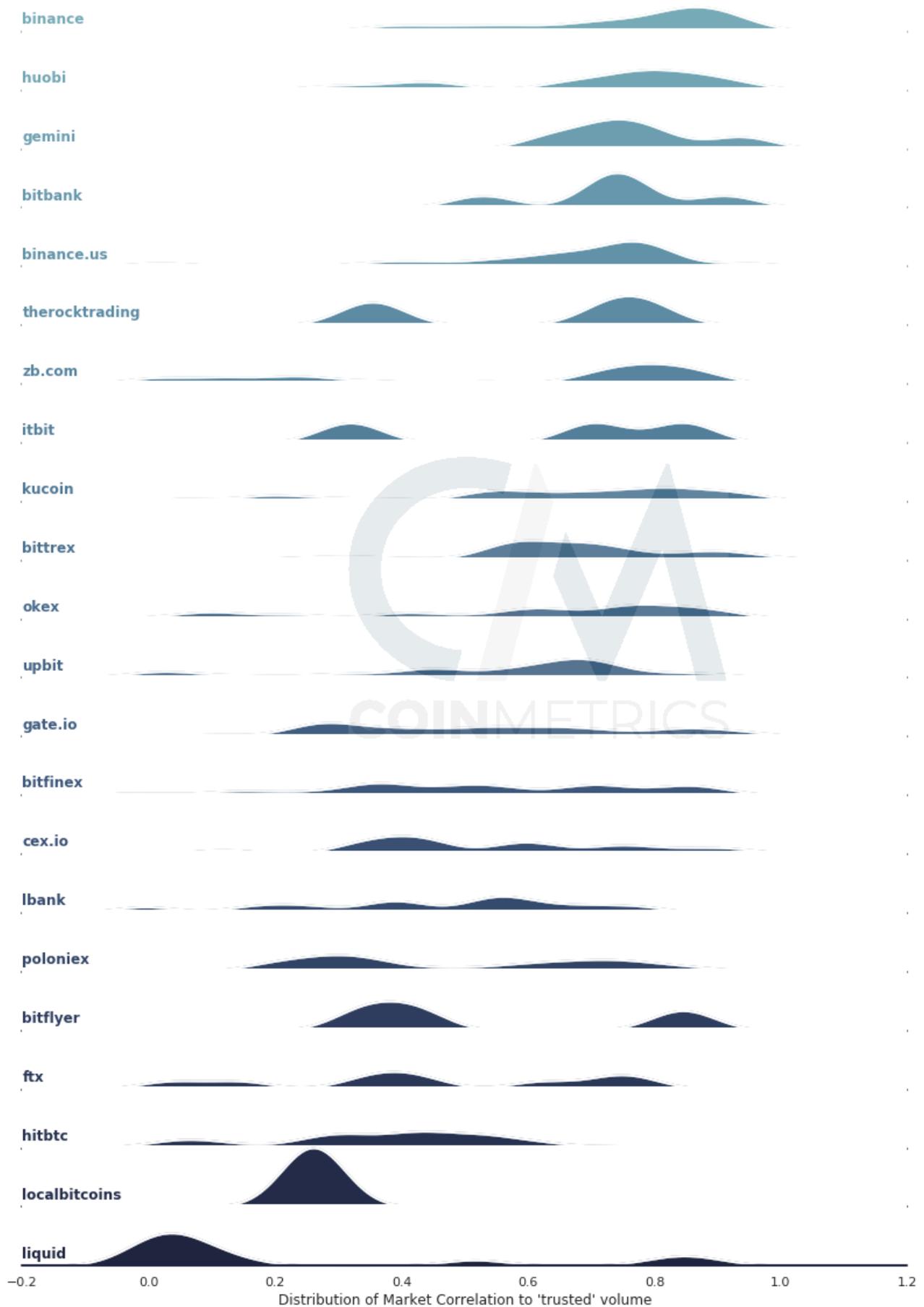


Above is a look at the correlations between the volume from exchanges and the volume from our 'trusted' control group for a few of the more well known assets, sorted by the correlation in the Bitcoin markets. We take the correlations (like those seen in the examples above) and look at them in their entirety to get a better sense for the overall dispersion of the exchange's markets' correlation to the 'trusted' control group.

Distribution of Market Correlation to Regulated Markets by Exchange - June 2020

Source: CM Market Data Feed





Above is a series of distributions with each row representing the dispersion of the correlation between an exchange's markets and the trusted markets. The further right the distribution, the more closely correlated it is with the trusted market's volume. The

further to the left, the less correlated. These markets are sorted by the median correlation of all of the exchange's qualifying markets. However, a median measure across an exchange's markets is not the best measure for economic activity.

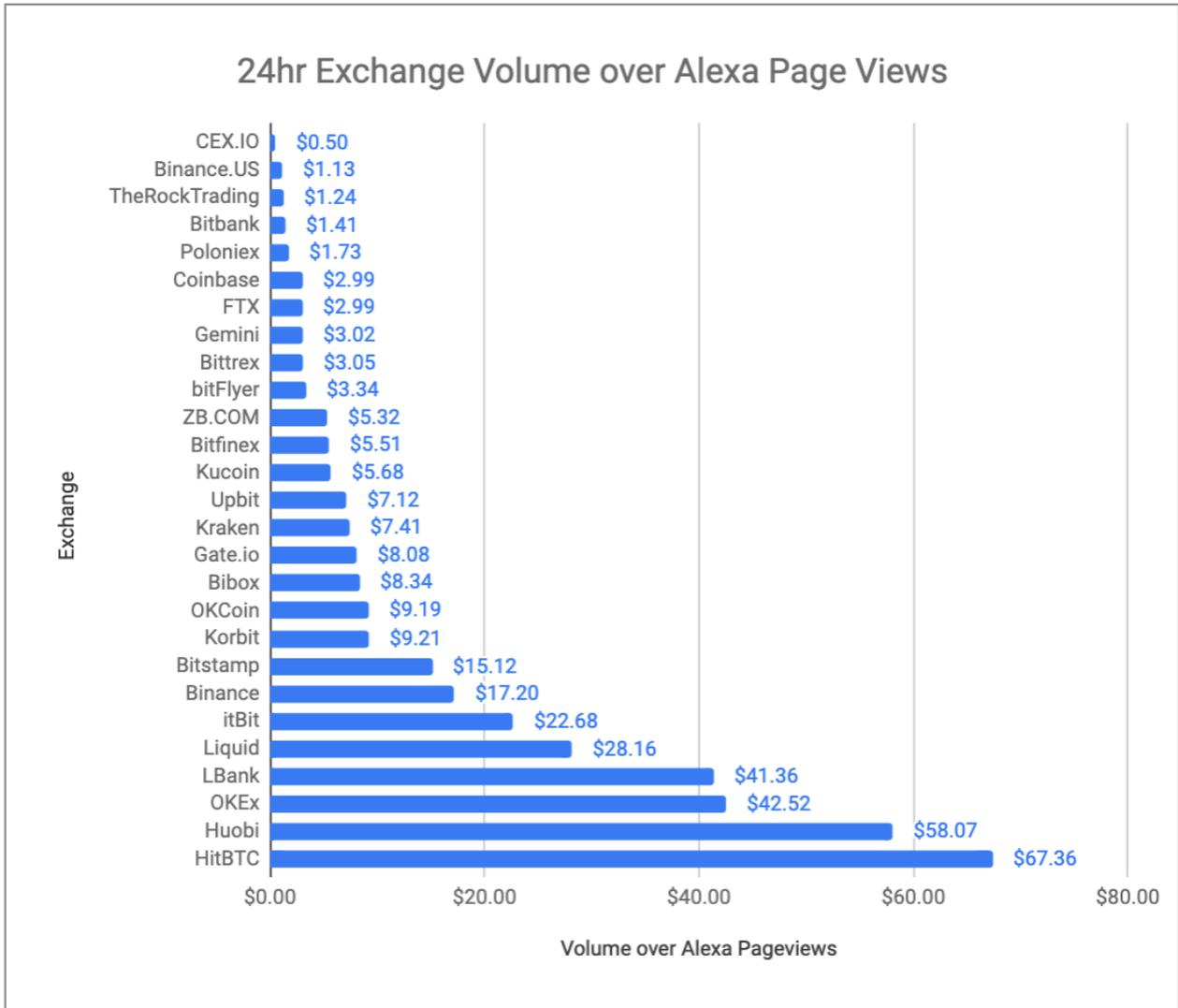
	Exchange	Volume Weighted Correlation	Rank
1	coinbase	98% 	1
2	kraken	96% 	2
3	bitstamp	96% 	3
4	binance	91% 	4
5	huobi	90% 	5
6	kucoin	90% 	6
7	gemini	88% 	7
8	cex.io	88% 	8
9	gate.io	87% 	9
10	liquid	87% 	10
11	bittrex	86% 	11
12	bitbank	86% 	12
13	bitfinex	83% 	13
14	bitflyer	83% 	14
15	itbit	82% 	15
16	binance.us	82% 	16
17	poloniex	80% 	17
18	zb.com	77% 	18
19	okex	76% 	19
20	upbit	76% 	20
21	therocktrading	75% 	21
22	ftx	70% 	22
23	hitbtc	61% 	23
24	lbank	44% 	24
25	localbitcoins	26% 	25

In order to make this measurement better represent the exchanges' overall volumes we created an aggregated volume weighted correlation based on the relative volumes of the exchanges' markets. This was accomplished by taking the volumes for the month of June 2020 and calculating the percentage that each base asset made up of the exchange's total qualified volume. Total qualified volume here is defined as volume in a base asset listed on two or more of the 'trusted' exchanges. We then used that percentage to create a volume weighted average of the correlations of each exchange. The result is the table above. We believe that the reasonable cutoff for this test is a volume weighted correlation to the 'trusted' markets of 80% or greater.

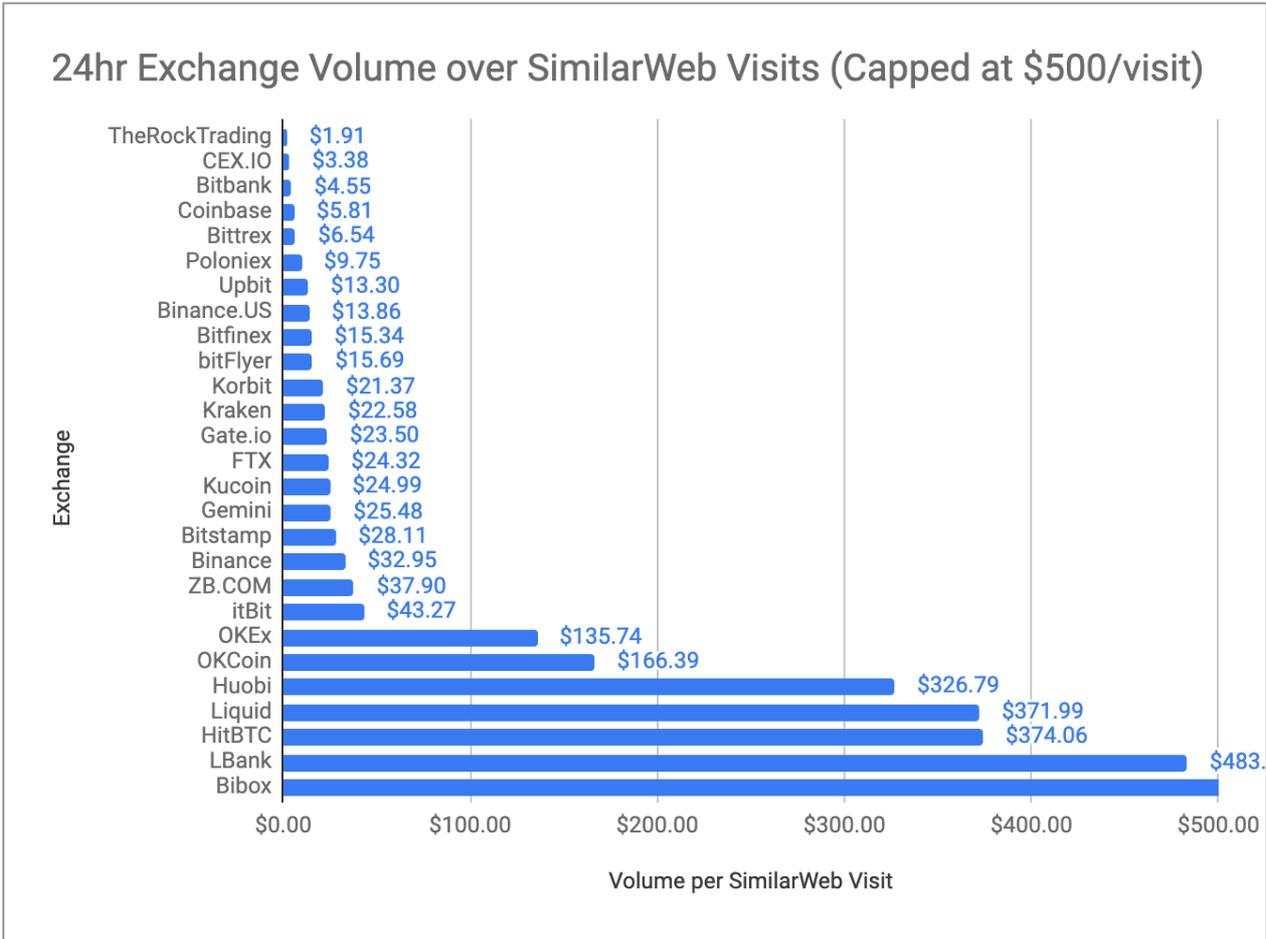
Part II: 24 Hour Volume Over Third Party Web Traffic Metrics

Another way of analyzing fake volume is to look at volume compared to the amount of users visiting an exchange. In theory, the amount of an exchange's volume should hold some relation to the amount of traders visiting their site. More users should lead to more volume and the inverse should hold as well. An exchange inflating volume numbers should tend to have a higher ratio of volume to traders relative to the other exchanges.

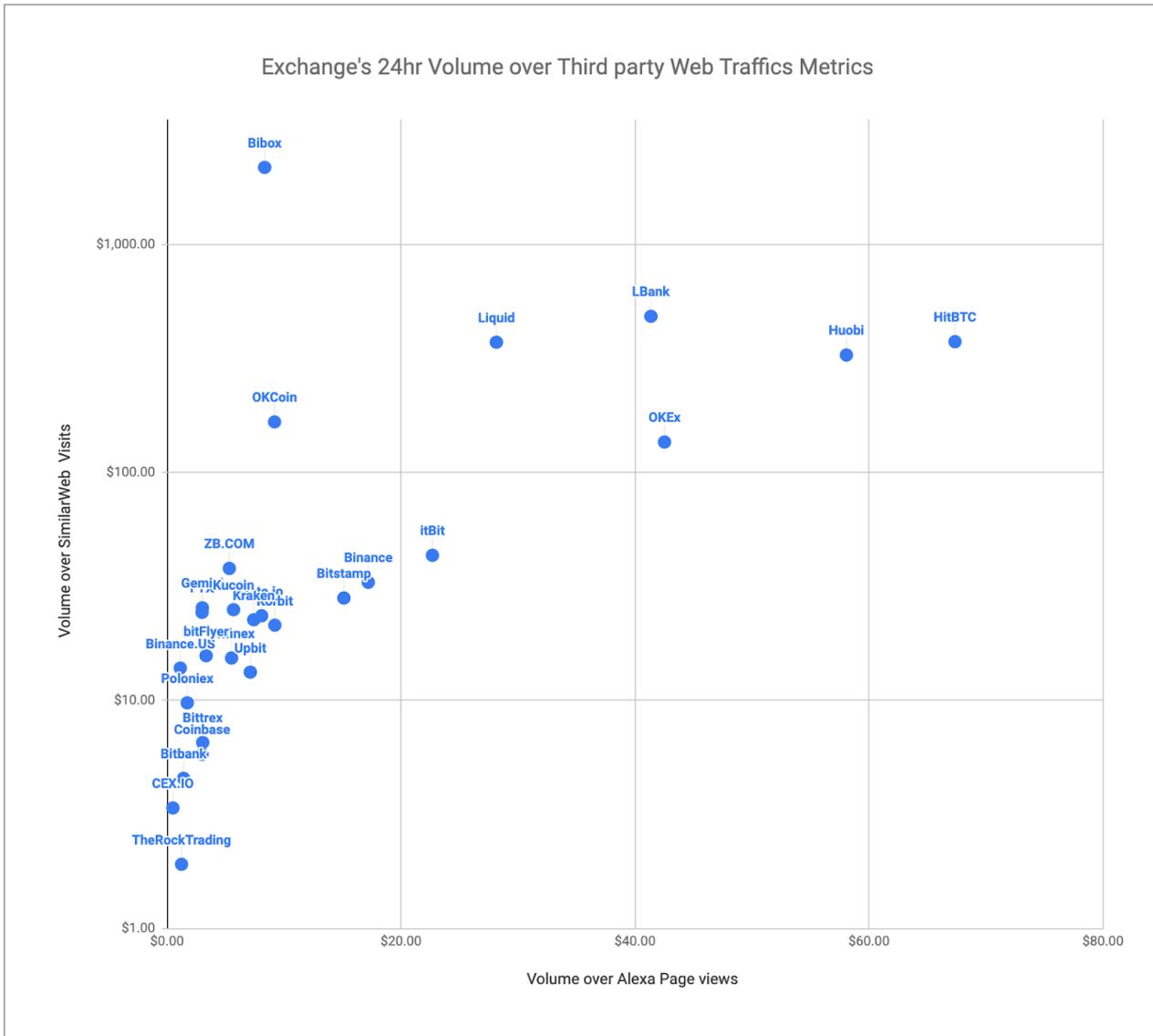
To test this hypothesis, we used third party data on web traffic from Alexa and SimilarWeb. Page visit metrics from these providers will be our proxy for exchange customers. We understand that this will not be perfect due to API traders that do not visit the webpage as actively. However, the values should still give us a reasonable directional indication of user activity.



Above is a look at the 24 hour volume for all spot exchanges in our coverage universe. This shows the volume divided by the Alexa page views per million. Exchanges with a relatively high ratio have a high amount of trading volume compared to the amount of users visiting the site. The regulated exchanges in our sample have an average ratio of \$9.05.



Here we do a similar comparison for SimilarWeb’s visits. Visits are a lower number than page views because you can have multiple pageviews per visit. It is our belief that this more accurately represents the volume inflation due to this distinction between unique visits vs page views. We have truncated the x-axis for visual purposes at \$500 per visit, however please note that the value for Bibox is \$2,502.67. Any exchanges with a ratio of greater than \$50 are relatively more suspicious based on the regulated exchange average of \$21.96.



Looking at both of those ratios together in the scatter plot above, you can begin to see outliers (please note that the y axis is in log scale). We believe that exchanges failing to meet both the thresholds set above (Volume to Alexa PageViews < \$20 and Volume to SimilarWeb Visits < \$100) fail this test.

It should be noted that these third-party web traffic metrics may have a western bias due to their techniques used to track traffic, discounting certain search engines or VPN use. You would expect to see two groupings of exchanges form here. The first being the 'retail' exchanges, those with a lot of customers making small transactions. On the plot above this is reasonably visible in the grouping around Coinbase, Gemini, Bittrex etc. The second group should be the exchanges where leverage is more common. Binance, OKEx would represent this exchange as they allow for traders to trade more per capita with the use of margin. The exchanges falling outside of the reasonable parameters exhibit a less consistent ratio of volume to visits/views than we are comfortable with.

Part III: Qualitative Factors

In addition to our quantitative analysis, we looked at a number of qualitative exchange features. We treated these as binary qualifications and assigned them each a weight

based on relative significance in determining trustworthiness, transparency, and authenticity. If a venue had the given feature, it was given full points. The following is the list of qualitative features and relative weights used in our analysis:

- a REST API 10
- a Websocket feed 5
- Historical trade data is available 10
- Real time trade data is available 10
- Real time order book data is available 10
- Has a Trading policy in place 2
- Has market surveillance 2
- Has regulatory oversight 5
- Requires KYC for basic trading 2
- Offers insurance for fiat holdings 2
- Offers insurance for crypto holdings 2
- Allows fiat deposits and withdrawals 5
- Markets quoted in fiat 2
- Trading fees 10
- Listing standards 2
- HQ is based in United States 5
- Does not have Trading competitions/contests 5

This features were chosen because they help demonstrate an exchanges transparency (e.g. API access, trade, orderbook data), fair market structure (e.g. trading rules, fiat access, KYC), and arm's length economic transactions (e.g. trading fees, no trading contests). Below is a ranking of the spot exchanges that we analyzed based on their total score from these factors.

Rank	Exchange	Qualitative Score
1	Coinbase	89
2	Binance.US	78
3	Gemini	77
4	Bitbank	76
5	itBit	75
6	Liquid	73
7	Kraken	71
8	CEX.IO	70
9	Bitstamp	68
10	TheRockTrading	67
11	bitFlyer	65
11	Bittrex	65
13	Binance	64
14	Bitfinex	63
15	Poloniex	62
16	Upbit	61
17	FTX	60
17	Gate.io	60
17	HitBTC	60
17	LBank	60
21	Korbit	58
22	Huobi	52
22	OKCoin	52
24	Bibox	50
24	Kucoin	50
24	OKEEx	50
24	ZB.COM	50

We decided on 50 and below as the cutoff point. Our rationale behind 50 points was that because this is a qualitative measure we should treat it fairly liberally and exclude only those in the bottom decile. Bibox, Kucoin, OKEEx and ZB.com were the only exchanges ruled out due to this metric.

Results

Coin Metrics - Trusted Volume Test Results - June 2020				
Exchange	Correlation	Qualitative	Web Traffic	Tests Passed
Bibox	PASS	FAIL	FAIL	1
Binance	PASS	PASS	PASS	3
Binance.US	PASS	PASS	PASS	3
Bitbank	PASS	PASS	PASS	3
Bitfinex	PASS	PASS	PASS	3
bitFlyer	PASS	PASS	PASS	3
Bitstamp	PASS	PASS	PASS	3
Bittrex	PASS	PASS	PASS	3
CEX.IO	PASS	PASS	PASS	3
Coinbase	PASS	PASS	PASS	3
FTX	FAIL	PASS	PASS	2
Gate.io	PASS	PASS	PASS	3
Gemini	PASS	PASS	PASS	3
HitBTC	FAIL	PASS	FAIL	1
Huobi	PASS	PASS	FAIL	2
itBit	PASS	PASS	PASS	3
Kraken	PASS	PASS	PASS	3
Kucoin	PASS	FAIL	PASS	2
LBank	FAIL	PASS	FAIL	1
Liquid	PASS	PASS	FAIL	2
OKCoin	PASS	PASS	FAIL	2
OKEEx	FAIL	FAIL	FAIL	0
Poloniex	PASS	PASS	PASS	3
TheRockTrading	FAIL	PASS	PASS	2
Upbit	FAIL	PASS	PASS	2
ZB.COM	FAIL	FAIL	PASS	1

Above are the outcomes for the correlation, web traffic, and qualitative metrics that we tested exchanges on. No exchange in our list failed all tests, likely due to the manual selection of the initial list of exchanges to test. Three exchanges only passed one test, eight passed two. The twelve exchanges that passed all three tests will be used for calculating what we believe to be the most accurate measure of daily spot volume.

These exchanges are:

- Binance
- Binance.US
- Bitbank
- Bitfinex

- bitFlyer
- Bitstamp
- Bittrex
- CEX.IO
- Coinbase
- Gate.io
- Gemini
- itBit
- Kraken
- Poloniex

This is the current passing group as of July 1, 2020. This list will be reviewed periodically and exchanges are likely to be added and removed overtime as their underlying data changes.

Note: certain exchanges, such as FTX have added a broader listing of spot assets since the end of June and subsequently improved performance on the volume correlation test.