Crypto Wash Trading*

Lin William Cong, Xi Li, Ke Tang, and Yang Yang

First draft: December 2019; current draft: June 2020

Abstract

To investigate whether centralized exchanges of cryptocurrencies engage in wash trading, we devise a set of statistical tests using first significant digits, size clustering, and tail distribution on transactions involving Bitcoin, Ethereum, etc., at 3 regulated and 26 unregulated major crypto exchanges. All regulated exchanges feature trades consistent with statistical benchmarks and behavioral patterns from traditional financial markets; most unregulated exchanges, especially those less prominent, exhibit anomalous patterns unexplained by investors' activities. We estimate that unregulated exchanges on average engage in wash trading for over 70% of their total reported volume. We also document that wash trades improve exchange ranking, and exchanges vying for dominance engage more in wash trading, consistent with their economic incentives. Finally, we discuss the regulatory implications of such manipulation in the emerging crypto sector and in financial markets in general.

Keywords: Bitcoin; Blockchain; Cryptocurrency; Forensic Finance; Regulation

JEL Classification: G18, G23, G29.

---

*Cong (will.cong@cornell.edu) is at the Cornell University Samuel Curtis Johnson Graduate School of Management; Li (xi.li@newcastle.ac.uk) is at Newcastle University Business School; Tang (ketang@tsinghua.edu.cn) and Yang (yangyang_tsu@mail.tsinghua.edu.cn) are at the Institute of Economics, School of Social Sciences, Tsinghua University.

1 Introduction

Since Bitcoin s inception in 2008, thousands of cryptocurrencies and blockchain projects have emerged.[1] The market capitalization of all cryptocurrencies peaked at $828 billion USD in 2017-2018 and is at $240 billion USD as of Jan 2020, with a total trading volume $8.8 trillion USD in the first quarter of 2020 alone (Helms, 2020). Meanwhile, both financial institutions and retail investors start to gain significant exposure to the cryptocurrency industry (Bogart, 2019; FCA, 2019; Fidelity, 2019; Henry, Huynh, & Nicholls, 2019).[2] As we write, crypto exchanges play a central role and are among the most profitable businesses in the ecosystem.

In the process vying for dominance in a lightly regulated market, some exchanges may gain advantage in ways that are ethically and legally questionable. In this paper, we investigate one manipulative behavior of crypto exchanges that entails faking trading in multiple major cryptocurrencies, including Bitcoin (BTC), Ether (ETH), Ripple (XRP), and Litecoin (LTC). Specifically, we introduce a series of tests based on well-established statistical principles and behavioral regularities to examine whether crypto exchanges follow empirical patterns observed in financial markets in general. By investigating the first significant digit distribution of trade size described by the Benford's Law, the clustering of trades surrounding round numbers, and the tail distribution of trade sizes described by the power law (Pareto-Levy law), we provide evidence that unregulated exchanges fake trading volume. Such "wash trading" on unregulated exchanges has significant impacts on their ranking and prominence within the industry, and likely impact the industry's long-term development.[3]

Wash trading in crypto exchanges warrants our attention for at least three reasons. First, crypto exchanges play a unique and essential role in the ecosystem. On the one hand, they

[1] Cryptocurrency, as a form of 'digital currency,' refers to digital tokens with the following characteristics: (i) they are typically native to a (decentralized or peer-to-peer) digital network or platform with liquidity secondary markets for trading; (ii) they are transferred and secured by modern cryptography; (iii) transaction data are often stored in a distributed database (with linked lists in the case of blockchain cryptocurrencies);with recordkeeping through consensus rules more decentralized than traditional systems. Cong, Li, and Wang (2019a, 2019b) and Cong, Li, and Xiao (2020) provide further institutional background.

[2] In a survey of 441 U.S. institutional investors, Fidelity claimed that 47% of respondents consider crypto-assets as the alternative asset class in their investment, and around 22% have already invested in crypto assets (Fidelity, 2019). The Financial Conduct Authority of the UK released a report in March 2019, showing that 25% of UK consumers could identify what a "cryptocurrency" is, and 3% had bought them (FCA, 2019). The Bank of Canada has conducted three iterations of the Bitcoin Omnibus Survey since 2016 to monitor trends in the adoption and usage of Bitcoin and other crypto-assets. They found that between 2016 and 2018, the percentage of Canadians who were aware of Bitcoin increased from 62% to 89%, and those who owned Bitcoin increased from 3% to 5% (Henry et al., 2019). Similar results were found by an online survey by The Harris Poll on behalf of Blockchain Capital (Bogart, 2019). The percentage of American adults that have heard of Bitcoin rose from 77% in October 2017, to 89% in April 2019, while the percentage of ownership has increased from 2% to 9%.

[3] Wash trading is, according to the U.S. Commodity Exchange Act, "Entering into, or purporting to enter into, transactions to give the appearance that purchases and sales have been made, without incurring market risk or changing the trader s market position." Definition from US Commodity Exchange Act can be found at https://www.cftc.gov/ConsumerProtection/EducationCenter/CFTCGlossary/glossary_wxyz.html

provide liquidity for various cryptocurrencies and facilitate price discovery like traditional exchanges. On the other hand, large crypto exchanges have expanded into upstream (e.g. mining) and downstream (e.g. payment) sectors, therefore wielding influence on all aspects of the industry. To some extent, crypto exchanges function as a complex of trading platforms, custodians, banks, and clearing houses. The regulatory vacuum led to hundreds of firms launching crypto exchanges and dozens of new entrants constantly emerge every year since 2011, due to low costs and barriers to entry. As such, crypto exchanges constitute a fitting starting point to understand the industry from both academic and regulatory perspectives.

Second, crypto exchanges have strong economic incentives to inflate trading volumes to manipulate the market. Trading volume is crucial for exchanges in general when it comes to marketing and advertising. Most crypto-exchanges generate profit by collection transaction fees and the profit crucially depends on the market share, with brand awareness and website traffic being two critical factors for customer acquisition, both of which heavily rely on public rankings in broadly recognized data tracking/ranking services or third-party websites or media (e.g., CoinMarketCap, CoinGecko, Bitcointalk, and Reddit). The most common ranking algorithm is based on the reported trading volume. Because exchanges offer more or less similar products and services, the competition is even fiercer than that in traditional markets.[4]

Third, wash trading has a strong negative impact on the development of the crypto market and on consumer welfare. The manipulative behavior artificially inflates the trading volume and falsifies liquidity in the market, leading to market participants' misjudging the trading environment and incurring potential losses. Wash trading is legally prohibited in the majority of financial markets (IOSCO, 2000). Even when present, wash traders tend to be institutional investors rather than exchanges that are heavily regulated. However, in the cryptocurrency industry, exchanges are the biggest suspects due to the lack of regulatory oversight. Consequently, wash trading damages the orderly development of the entire industry by hindering price discovery and crowding out regular trades --- a variant of "Gresham's law."

[4] Unlike established brands with user stickiness and network effect (Halaburda & Gandal, 2016; Cong, Miao, Tang, & Xie, 2019), newcomers (with little reputation) are more tempted to pursue high rankings that might be achieved via wash trading. Top ranked exchanges are thus not necessarily reputable and secure and investors who are misled to them could face substantial risks. For example, Fcoin, which become insolvent in February 2020, previously ranked 56th on CoinMarketCap. However, Gemini, a crypto-exchanged certified and regulated by the New York State Department of Finance, is listed 124th on the second page of CoinMarketCap. https://coinmarketcap.com/rankings/exchanges/reported/2/ accessed 2019/12/29

In the analyses, we collect cryptocurrency transactions of 29 major exchanges from the unique database of "TokenInsight." [5] The dataset consists of information for every transaction from 00:00 July 09th to 23:59 November 03rd, 2019. Our data contains information from crypto exchanges such as trading volume, reputation, firm size, and age. For each exchange, we examine trading of four major cryptocurrencies against US dollars (USD) — Bitcoin (BTC), Ethereum (ETH), Litecoin (LTC) and Ripple (XRP) — which are most widely recognized and heavily traded in the cryptocurrency market.[6]

We define regulated exchanges by adopting the regulatory definitions from the state of New York, which has one of the earliest and most comprehensive regulatory frameworks in the world. Regulated exchanges are issued BitLicense and regulated by the New York State Department of Financial Services. Unregulated exchanges are further divided into unregulated Tier-1 and Tier-2 exchanges based on reputation and popularity: unregulated Tier-1 exchanges refer to those unregulated exchanges which are renowned and ranked at the top 800 in the finance/investment section of SimilarWeb[7] ; the rest are unregulated Tier-2 exchanges.

We provide empirical evidence that wash trading broadly exists in unregulated exchanges. To start, we examine the first significant digit for each transaction size and check the frequency distribution of leading digits against Benford's law for each exchange. Benford's law represents a statistical benchmark that occurs naturally in sciences and social sciences, and is widely used to detect frauds in macroeconomic, accounting and engineering fields (e.g., (Durtschi, Hillison, & Pacini, 2004; Z. Li, Cong, & Wang, 2004). We find the first-digit distributions for all regulated exchanges satisfy Benford's law, but those for unregulated exchanges deviate significantly, implying that transaction data are not naturally generated from actual trading.

We next exploit a behavioral regularity in trading: clustering around certain transaction sizes. Rounded numbers are routinely used as cognitive reference points in individuals' decision-making. For example, multiples of 10 are commonly regarded as cognitive reference points in the decimal system (Rosch, 1975). This round-number bias is broadly observed in finance (Chen, 2018; Hervé & Schwienbacher, 2018; Kandel, Sarig, & Wohl, 2001; Kuo, Lin, & Zhao, 2015; Mitchell, 2001) and even analysts' forecasts (Clarkson, Nekrasov, Simon, &

[5] TokenInsight is a third-party data provider who offers consulting, rating and research reports for cryptocurrency-related business.
[6] Since US dollars (USD) are only allowed to exchange in three US regulated exchanges (R1, R2 and R3), digital dollars (e.g. Tether-symbol USDT, which is designed to be pegged to the US dollar) are commonly used as substitutes and widely accepted by majority of trading platforms. Therefore, we treat cryptocurrency-USD pairs and cryptocurrency-USDT pairs the same in our study.
[7] https://www.similarweb.com

Tutticci, 2015; Roger, Roger, & Schatt, 2018). Most cryptocurrencies can be traded at fractions of a unit. If the amount worth of $1 is set as a base unit of mental account, we expect that trades concentrate around multiples of 100, 500, 1000, 5000 and 10,000 base units for each cryptocurrency.[8] Authentic transactions by investors should exhibit size clustering effects. We examine the presence of size clustering effects using t-test to compare the trade frequencies between trades at and around rounded size. We find significant clustering on regulated exchanges whereas the trades on unregulated exchanges, especially Tier-2 exchanges, exhibit little clustering.

Our third test explores the distribution of observed trade size, which has heavy tails characterized by power law in traditional financial markets and is the common property of empirical regularities in economics (Gabaix, Gopikrishnan, Plerou, & Stanley, 2003a).[9] We fit a power-law distribution and estimate the exponent parameter, in addition to graphically inspecting the tail distributions on log-log scale. Regulated exchanges exhibit Pareto–Lévy tails in all cryptocurrency trades and the scaling parameters lie between 1 and 2, consistent with regularity in traditional financial markets. 50% of unregulated Tier-1 exchanges show inconsistency with power law (Pareto–Lévy) in at least one cryptocurrency; 75% of Tier-2 exchanges fail to display power law (Pareto–Lévy) decay in trade distribution of any cryptocurrency. We conclude that unregulated exchanges, especially Tier-2 exchanges, produce suspicious trading patterns according to their trade size distributions.

Finally, we estimate the fraction of fake volume on unregulated exchanges using a metric derived from the clustering regularity. To avoid being easily detected and achieve good efficiency, exchanges naturally would prefer to use machine-generated fake orders and make these orders not be too big in size. Therefore, wash trades, generated by automated programs, are very likely to have a low level of roundness, i.e. the effective number of decimals for trades is big. Human transactions have a higher level of roundness than artificial ones; for example, a man-made order in the size of 0.1 BTC has a higher level of roundness than the machine-made 1.14357 BTC. It is possible that valid algorithmic trading exists in legitimate exchanges and human trades can be unrounded due to special needs, we thus adopt a benchmark ratio of unrounded trades to rounded human trades based on calculations from the regulated exchanges. The extra unrounded trades above the ratio are thus treated as an estimation of wash trades on unregulated exchanges. Following this

[8] For example, value of 10^{-4} BTC approaches $1. 10^{-4} BTC is our base unit of BTC. We check the clustering effect at 0.01, 0.05, 0.1, 0.5 and 1 BTC. Similarly, 0.001 ETH is the unit of ETH, and clustering are checked at 0.1, 0.5, 1, 5 and 10ETH. The base units of LTC and XRP are 0.01 LTC and 1 XRP, respectively.

[9] For example, power law decay is observed in returns, trade size and volume in equity markets (Gopikrishnan, Plerou, Amaral, Meyer, & Stanley, 1999; Gopikrishnan, Plerou, Gabaix, & Stanley, 2000; Plerou, Gopikrishnan, Amaral, Gabaix, & Stanley, 2000; Plerou & Stanley, 2007). Recently, M.-Y. Li, Cai, Gu, and Zhou (2019) and Schnaubelt, Rende, and Krauss (2019) document that the trade-size distribution for bitcoin follows a power law tail.

principle, we estimate a wash trading volume to be around 70% of the total trading volume on the unregulated exchanges on average. In particular, the wash trades on the twelve Tier-2 exchanges are estimated to be more than 80% of the total trade volume. Our estimates combined with the reported volumes in Helms (2020) translate into wash trading of over $4.5 Trillion USD in spot markets and over 1.5 Trillion USD in derivatives markets in the first quarter of 2020 alone.

While there is no evidence that wash trading affects cryptocurrency prices directly, we obtain proprietary data on historical ranking and trading volume information from CoinMarketCap and show that exchange ranking is highly correlated with (log) trading volumes: 10% increase of trading volume moves ranks up by 2 places. Therefore, when wash trading constitutes 70% of the total trading volume, it can move an exchange's rank by 46 positions. We also find that top-ranked exchanges engage in more wash trading in terms of percentage of their total volume in future, creating a mutual reinforcement between ranking and wash trading. The finding is intuitive — wash traders benefit from being ranked highly because they likely garner greater publicity, user base, and revenue from collecting transaction fees. Because liquidity begets liquidity and the winner-takes-all nature of the competition, exchanges vying for the top positions could be especially enticed to wash trade.

Overall, we show that wash trading appears prevalent on unregulated cryptocurrency exchanges, a phenomenon fuelled by the current business models of exchanges and ranking systems. Regulated crypto exchanges, in contrast, have committed considerable amount of resources in compliance and license acquisition, and therefore do not wash trade, anticipating severe punishment including financial loss, discredit and legal consequences of any violations. Moreover, to discourage wash trading behavior, the current third-party ranking companies should think alternative ways other than trading volumes in ranking exchanges. Our study is the first systematic demonstration of the effect of regulation in the cryptocurrency markets, which has implications for investor protection and financial stability. We, therefore, offer a concrete starting point for government regulation and third-party supervision in the crypto market by not only convincingly exposing wash trading of exchanges, but also providing a set of tools to identify and combat non-compliant and unethical behaves in general.

We contribute to the recent studies on cryptocurrencies in several ways. To the best of our knowledge, our paper is the first academic study to systematically and rigorously analyze crypto wash trading as an industry-wide phenomenon, beyond anecdotal evidence or case study in the media. As more and more institutional and retail investors start trading cryptocurrencies, wash trading has become a concern too big to ignore. We also speak to the

debates on blockchain centralization (e.g., Cong, He, & Li, 2020) by highlighting a detriment of over-concentration of the operation scope of crypto exchanges.

Our paper relates to the literature on manipulation in cryptocurrency markets.[10] For example, Gandal, Hamrick, Moore, and Oberman (2018) and Griffin and Shams (2018a) study manipulative behavior in Bitcoin and Tether; Foley, Karlsen, and Putninš (2019) study the illegal usage of cryptocurrencies; studies such as Li, Shin, and Wang (2019) document pump-and-dump patterns in various cryptocurrencies. However, their data do not allow them to examine wash trading on various crypto exchanges. Our unique and comprehensive data set allows us to apply scientific yet straightforward procedures to detect crypto wash trading unambiguously. Moreover, unlike market manipulations documented in Gandal et al. (2018) and Griffin and Shams (2018a), exchanges engage in wash trading not to manipulate a cryptocurrency's underlying price, but to increase trading volume and transaction fees to generate greater profits for the exchanges themselves.

Wash trading distorts price, volume, and volatility (Aggarwal & Wu, 2006; Cumming, Johan, & Li, 2011). The manipulation harms the efficiency of financial markets (Aggarwal & Wu, 2006), negatively affects the investors' confidence, and deters the participation of new customers (Imisiker & Tas, 2018). Therefore, it is banned worldwide, and mainly exists in less efficient markets, such as over-the-counter markets and emerging financial markets. Most of the academic literature on wash trading focuses on investor behavior. We add by investigating wash trading at the exchange level with evidence from the emerging crypto markets, which provide unique cases in modern time to allow us to detect and analyze the manipulation engaged by more informed participants.

The paper proceeds as follows. Section 2 introduces the history, current development, and regulatory status of the cryptocurrency exchanges. Section 3 describes the data and stylized patterns. Section 4 presents the methodologies of wash-trading detection and report our empirical findings. Section 5 concludes.

[10] Our paper also contributes to the general literature on trade-based manipulation and forensic finance — the use of economic and financial knowledge to discover or substantiate evidence of criminal wrongdoing that meets standards in a court of law. See Ritter (2008) and Zitzewitz (2012) for surveys of the field. Forensic finance is closely associated with forensic accounting, of which there is significant overlap, which focuses on using accounting and auditing standards and techniques to analyse an organization's financial statements to discover evidence of wrongdoing. The empirical literature on general market manipulations is growing (Christie & Schultz, 1994; Griffin & Shams, 2018a, 2018b) with well-founded theoretical foundations (Allen & Gale, 1992; Jarrow, 1992).

2 Institutional Background of Crypto Exchanges: Development and Regulation

We introduce the institutional background of crypto exchanges. Readers familiar with the cryptocurrency market and industry may skip this Section.

Satoshi Nakamoto introduced Bitcoin in October 2008 and launched it three months later with the headline in the Times embedded in the genesis block.[11] Because Bitcoin software was developed in an open-source environment, other "altcoins" (referring to 'alternative coins to Bitcoin') quickly emerged to imitate and improve on the first few cryptocurrencies. For example, Ethereum, EOS, and Tron were developed as public platforms for smart contracts and decentralized applications, but also come with their native cryptocurrencies on their own blockchains.[12] Nowadays, over 4000 cryptocurrencies have been launched and circulated globally. The total market capitalization of all cryptocurrencies grew from non-existent to a peak value of $828 billion in January 2018, larger than Visa and Facebook combined.[13]

Increasing awareness and adoption among financial institutions and retail investors shape a flourishing and increasingly sophisticated ecosystem involving mining, payment companies, wallets, dapp (decentralized application), and crypto exchange (Hileman & Rauchs, 2017). Crypto-exchanges, centralized gateways that facilitate money flow between fiat currency and (decentralized) cryptocurrency systems, play a critical and dominant role in the industry (Griffin & Shams, 2018a). To date, over 300 exchanges provide cryptocurrency services around the globe, often with leverage facilities and derivatives on cryptocurrencies. Incumbents exit and new competitors keep emerging due to loose regulatory standards.

Currently, the cryptocurrency trading volume on exchanges (likely mostly speculation activities) is much higher than the on-chain transaction volume (likely actual usage). With considerable traffic, exchanges usually hold a large number of various cryptocurrencies because of liquidity demand and custody for customers. As a result, they wield enormous

[11] "The Times 03/Jan/2009 Chancellor on brink of second bailout for banks."

[12] Monero, Zcash, and Dash were created to address Bitcoin's privacy limitations and shortcomings. Other cryptocurrencies focused on applications content creation and copyright (Steem, Ink), on social/communication (KEY, SNT), on the internet of things (IOTA, QTUM) and computation power/cloud storage (SC, FCT), among many others.

[13] Cryptocurrency total market cap peaked in Jan 7, 2018. From ycharts.com, Visa (Jan 3, 2018 $268.10), Facebook (Jan 3, 2018 $536.65B)

power in the industry. This is somewhat ironic, given the initial ideas of decentralization and financial democratization for the industry.[14]

In early days, regulators deemed the cryptocurrency industry size-limited and unimportant, hence were reluctant to regulate crypto industry. It was widely believed that all crypto-exchanges had, to some extent, engaged in non-compliant and unethical behavior (Gandal et al., 2018; Moore & Christin, 2013; Moore et al., 2018), which might later on threaten global financial stability as the crypto industry is getting larger and larger. For instance, before the trading process, customers are often required by exchanges to submit official ID documents for KYC (know your customer) check. This leads to the risk of identity information leaking when the exchanges' database is breached. After successful registration of an exchange account, users often find it challenging to channel fiat money into or withdraw out of exchanges. This tip of the iceberg reveals that customers may be exposed to various operational risks, credit risk, and settlement risk.

Exchanges usually hold substantial funds from users' accounts (both in fiat and cryptocurrencies) without proper custody and insurance, which raises severe concerns. Moore and Christin (2013) and Moore et al. (2018) examine the failure of Bitcoin exchanges from 2010 to 2015, and show that security breaches greatly increase the probability of exchange closure. The cryptocurrency ecosystem has already seen shortcomings in asset storage and numerous cyber-attacks bring down numerous exchanges (even previously dominant exchanges such as Mt.Gox), which in turn affects their unprotected customers. Most often, implied counterparty risk manifests in the form of notorious 'runaway bosses' incidents or exit scams (malicious closure of exchanges and stealing users' funds). For example, the once No.1 transaction-mining exchange Fcoin claimed insolvent all of a sudden with $130 million client's funds missing (Zhao, 2020). Some exchanges are quagmire in Ponzi schemes, in which case existing scam is not beyond expectation. For instance Xcoinx, which is operated by the Onecoin team.[15] Others include Coinroom (Alexandre, 2019), Cobinhood (Palmer, 2020), OKUEX, Soxex, etc. The list is expanding and not likely to stop.

Profit-driven exchanges may also take advantage of information asymmetry or even directly act against users interests through various market manipulation measures. In an unregulated environment, unethical cryptocurrency exchange can be "both a referee and a player" at the same time. Gandal et al. (2018) investigate the manipulative trading in Mt.Gox, a Bitcoin exchange, over the period from February to November 2013, and find a

[14] While pioneers build decentralized exchanges, matching buyers and sellers directly through smart contracts and mitigate risks related to middlemen, they constitute a small portion of exchanges in crypto industry and are not our focus.
[15] Onecoin, Wikipedia. https://en.wikipedia.org/wiki/OneCoin

suspicious trader called “Markus”, likely to be an exchange owned account, participates in manipulative trading. Our paper also shows that many exchanges have engaged in wash trading, likely aiming to improve their ranking or attract more customers.

The generally lack of consumer protection in the cryptocurrency industry aggravates the situation. Consumers’ legitimate rights and interests heavily rely on exchanges’ self-discipline and good faith. If user interests are undermined in incidents such as hacking or bankruptcy, victims get little compensation from either exchanges or third-party insurance companies.

As such, risks in the cryptocurrency exchange ecosystem have drawn significant attention from regulatory authorities in recent years. Regulatory authorities in multiple jurisdictions have published statements to warn the public about the risks (Yu, 2018), and have built internal divisions and created new institutions to closely monitor the development of the cryptocurrency industry (Brett, 2019). A number of authorities (e.g. Bank of Canada, UK Financial Conduct Authority, New York Federal Reserve Bank) have conducted surveys to investigate the awareness and adoption of cryptocurrency in retail and institutional investors. Since 2017, official cryptocurrency documentation and guidelines have been released by regulatory agencies in around 20 countries and territories, including United States, European Union, United Kingdom, China, Japan, etc. (Blandin et al., 2019).

However, some unique features of the cryptocurrency industry render these attempts futile and ineffective. For one, regulatory frameworks are designed and developed differently from countries to countries. Currently, there is no consensus on the correct regulatory approach; it is unlikely that a universal scheme would emerge soon to properly address the scope of the industry due to the nature of the cryptocurrency, namely, anonymity, multi-nationality and high-technicality. The intention and infrastructure for sharing information and collaborative effort are also lacking. Overall, new regulatory tools and a well-integrated regulatory framework are in dire need.

3 Data

Data used in this research come from multiple resources. The cryptocurrency trading history data is obtained from the TokenInsight database, which provides ratings and research reports about the cryptocurrency industry as an independent third-party. The transaction data in this database is fetched through each exchange’s official API (Application Programming Interface). Our dataset contains the whole reported trade history of 29 major exchanges over three months from 00:00:00 July 09th to 23:59:59 November 03rd, 2019. Each

record of the transaction contains the exchange information, unique transaction ID, timestamp, price, amount of cryptocurrency traded, and trade pair symbol.

Exchanges-related data is collected from both their official websites and various data tracking and analysis platforms. We gather rankings data from SimilarWeb,[16] Alexa,[17] and CoinMarketCap.[18] SimilarWeb and Alexa are online platforms that track and analyze website popularity and providing rankings by web traffic. CoinMarketCap is a widely acknowledged data provider in the cryptocurrency industry, where we obtain trading volumes and ranks of about 300 exchanges during the sample period.

In this study, crypto exchanges are classified as either regulated or unregulated. The regulation entity of New York State, the New York State Department of Financial Services (NYSDFS), is one of the fastest agencies in establishing regulation over cryptocurrencies. It has developed a world-leading regulatory framework for the cryptocurrency industry. Hence, we categorize three exchanges (R1, R2, and R3) issued with the BitLicense by NYSDFS as regulated exchanges as all the three operate under the supervision of NYSDFS. Other 26 non-compliant exchanges are classified as unregulated and are further divided into 10 Tier-1 unregulated and 16 Tier-2 unregulated exchanges. Compared with Tier-2 unregulated exchanges, Tier-1 unregulated exchanges are ranked in the top 800 based on "SimilarWeb" website traffic ranking of the investment category during the sample period. We consider brand awareness and web traffic as the main factors in our classification since these two factors play an essential role in exchanges regarding customer acquisition and competition. Therefore, intuitively, unregulated Tier-1 exchanges have a better reputation than Tier-2 exchanges.

[Insert Table 1]

We summarize characteristics of exchanges in Table 1, including age, trading volume and ranks from different metrics. Note that ages for exchanges refer to the period from their dates of the establishment to December 2019. In Table 1, all of the regulated exchanges have survived for at least five years to date. However, most of the unregulated Tier-2 exchanges were launched in 2017 and 2018, while Tier-1 exchanges are generally older. This finding indicates that exchanges benefit from the long-term operation. Trade volume shows little correlation with our classification on exchanges: Some unregulated exchanges have much larger trading volumes compared with regulated exchanges. For example, U4, an unregulated Tier-2 exchange, has $50,944 million volume while R2's volume is only $15,212 million.

[16] Ranking is based on report over the period from Aug 2019 to Oct 2019 https://www.similarweb.com/.
[17] Ranking is accessed through https://www.alexa.com/siteinfo in Nov/15/2019.
[18] Ranking is based on daily trade volume reported by CoinmarketCap https://coinmarketcap.com/ in Jan/04/2020.

The trading volume of different unregulated exchanges varies significantly. U9 has only dozens of millions, while a large portion of unregulated exchanges exceed tens of billions in the sample period. When looking at the ranking of webpage traffic, we find regulated exchanges, especially R1 and R3, fall behind many unregulated Tier-1 exchanges; while R2 has the highest trading volume among regulated exchanges and a better rank under both ranking algorithms. CoinMarketCap ranks exchanges by trading volumes; seven unregulated Tier-2 exchanges rank Top 20 and outperform the majority of unregulated Tier-1 and regulated exchanges. Although trading-volume ranks cannot fully represent the quality of exchanges, it is applied by most ranking agencies. Thus, cryptocurrency investors are likely to choose an exchange based on these trading-volume based ranks. Therefore, unregulated exchanges, especially ones that are launched later, are motivated to engage in wash trading in order to achieve higher rankings and additional customer acquisitions given no regulations imposed for the time being. The next section presents formal empirical results on the detection of wash trading for different exchanges.

4 Empirical Tests and Quantification for Wash Trading

We present empirical evidence of wash trading in crypto-exchanges, inside the trading activities of 4 major trading pairs (BTC/USD, ETH/USD, LTC/USD and XRP/USD).[19] In particular, we examine the properties of *trade sizes* of each crypto exchange and devise a set of statistical tests based on the Benford's law, round-number clustering and the power law. Finally, we estimate fake volume for unregulated exchanges to quantify the extent of wash trading.

4.1 Distribution of First Significant Digit

To detect wash trades, we collect first-digits of every cryptocurrency transaction[20] on every exchange. We then investigate whether the first-digit distributions conform to the pattern implied by the Benford's law. Inconsistency with Benford's law suggests manipulation in trade sizes and therefore implies wash trading.

4.1.1 Introduction to Benford's Law

Benford's law or Newcomb–Benford law was first proposed by an American astronomer Simon Newcomb in 1881 inspired by the degree of abrasion in different parts of books in a library. Though initially unnoticed, the proposed law was rediscovered and elaborated in detail by an American physicist Frank Benford (1938). It describes a first-digit distribution,

[19] For brevity, results of all crypto-exchanges are reported in Appendix A. LTC/USD data is not available in unregulated exchange UT7, U1, U6 and U9. XRP/USD data is not available in regulated exchange R3 and unregulated exchanges U1 and U6.
[20] Trade size is denoted as the amount of cryptocurrency traded in each transaction.

which indicates that smaller digits (e.g., 1, 2) appeared more frequently in leading positions than larger digits (e.g., 7, 8, and 9) in various *natural* datasets. Moreover, Benford's law suggests a non-uniform distribution of the nine possible digits. Instead, the value of a digit (from 1 to 9) is negatively associated with the probability that the digit occurs in the leading digit of a dataset. The probability of number N appears as the first significant digit follows the distribution (Benford, 1938):

$$\text{Prob}(N\ is\ the\ first\ significant\ digit) = \log_{10}(1 + N^{-1})\ ,\quad N \in \{1,2,3,4,5,6,7,8,9\} \tag{1}$$

Specifically, the probability of digit 1 found in the first effective position is 30.10%. Digit 2 and 3 have a probability of 17.60% and 12.50%, respectively. The probability of the rest (from 4 to 9) as initial digit is less than 10% and decrease as the value of the first digit increase.[21] Previous statistical literature provides mathematical evidence to illustrate the presence of Benford's law (e.g., (T. P. Hill, 1995, 1998; Pinkham, 1961).

Naturally, the Benford's law holds in the datasets with stochastic settings, in which data are randomly and independently chosen from one distribution or mixed random sampling from various distributions. Apart from natural or sequential data (e.g., mobile numbers), deterministic samples with exponential growth or decay also follow the Benford's law when numbers are expressed in base 10. Benford's law has been effectively applied to test the reliability of data and diagnose the presence of manipulation or irregularities in datasets. For instance, Sambridge, Tkalčić, and Jackson (2010) presented that Benford s law holds 15 sets of modern observations drawn from the fields of physics, astronomy, geophysics, chemistry, engineering and mathematics. In economics, Benford's law is introduced to detect and identify fraudulent data in areas such as tax payments, accounting and finance (Durtschi et al., 2004; Nigrini, 1996). Günnel and Tödter (2009) find that the data published in economic journals and forecasting is broadly distributed according to Benford's law. Gonzalez-Garcia (2009) apply Benford's law to test the quality of macroeconomic data.

4.1.2 Detecting Violation of the Benford's Law

We check whether the leading digits of trade sizes follow the Benford's law (as shown in Equation 1) for the 29 exchanges. Figure 1 depicts the first-digit distribution for four cryptocurrencies of representative regulated exchanges and two suspicious unregulated exchanges, the rest of unregulated exchanges are shown in Appendix A. Bars show the proportion of transactions in which the trade size has the leading-digit as integer *i*. Dots represent the frequency distribution implied by Benford's law.

[21] The probability of digit 4, 5, 6, 7, 8, 9 in the first effective position is 9.7%, 7.9%, 6.7%, 5.8%, 5.1% and 4.6%, respectively.

[Insert Figure 1]

In general, first-digit patterns of all regulated exchanges comply with the Benford's law regardless of types of cryptocurrencies. For instance, in R2 exchange (see Figure 1), 32.75% of BTC trades and 32.73% of ETH trades have "1" as the leading digit, consistent with the benchmark frequency of 30.10% in Benford's law. In unregulated exchanges, including Tier-1 and Tier-2, around 50 % of exchanges show an apparent discrepancy with Benford's Law in at least one type of cryptocurrency, with a more serious violation for Tier-2 exchanges. Disconformity with the Benford's Law is observed in nine unregulated Tier-2 exchanges, among them, seven violates in at least two cryptocurrencies.

For example, as shown in Figure 1, all cryptocurrencies in U8 violate the Benford's law: 55.45% of LTC trades and 64.65% of XRP trades that start with the digit "1" as a comparison to 30.1% in the Benford's law; while trades with "5" as the first digit account for 36.53% and 40% in BTC and ETH transactions, much more than the 7.9% in the Benford's Law. Figure 1 also shows that the frequency of number "5" as the leading digit in UT3 outweighs 7.9% predicted by Benford's law in BTC trades and XRP trades, respectively. UT6 (Appendix A Panel UT6) has a 10 % and 7.7% probability, respectively for BTC and ETH, more than the probability in Benford's law with "1" as the first digit. UT7 (Appendix A Panel UT7) shows an irregular distribution of the first digit in BTC and ETH trade sizes, different from the monotonic decreasing pattern in the Benford's law. In U5 (see Appendix A U5), 61.11% of BTC transactions begin with the digit "1" while for LTC it is only 15.27%, all heavily deviated from the benchmark 30.1% in the Benford's law. A similar issue is observed in U2, U3, and U7, where trades concentrated at abnormal initial digits. In U9, U10, and U14, the trade size distribution of the leading digit is not monotonically decreasing.

[Insert Table 2]

We employ the Chi-squared test to quantitatively assess whether first-digit distributions conform with Benford's law (see Table 2). Firstly, trades of regulated exchanges follow the first-digit distribution predicted by Benford's law. Secondly, most of the unregulated Tier-1 exchanges are in a similar fashion as regulated ones. However, digit patterns of UT3 are inconsistent with the Benford's law in BTC and XRP trades, with a significance level of 1%. Moreover, five Tier-2 exchanges (U5, U7, U8, U9 and U14) have prominent divergence from Benford's law in most cryptocurrencies. Other unregulated exchanges show significant differences in some cryptocurrencies. For example, UT7 and UT9 violate the Benford's law in BTC and XRP in a 1% confidence level; U2 and U10 fail in BTC and XRP at a 1% confidence level; U2 and U3 fail at a 5% confidence level in ETH.

To conclude, this section documents that all regulated exchanges show consistency with the Benford's law. 20% of unregulated Tier-1 exchanges violate the Benford's law in at least one cryptocurrency, at 5% significance level. 50% of Tier-2 exchanges fail to follow Benford's law in at least one cryptocurrency. These findings overall suggest the existence of artificial and suspicious wash trades in unregulated exchanges.

4.2 Trade Size Clustering

In this section, we analyze the trade size pattern and investigate whether the trades in crypto exchanges demonstrate retail traders' behavioral regularity.

In traditional financial markets, the psychological and behavioral bias of humans heavily affects financial decision-making, therefore it leads to human-specific features in trade patterns. The clustering effect is one of the typical human behavioral regularities. It widely reflects traders' tendencies for rounded trade sizes and rounded prices in traditional financial markets. For instance, Alexander and Peterson (2007) show that in the New York Stock Exchange (NYSE) and Nasdaq, higher proportions of trades occur round sizes that are multiples of 500, 1000 or 5000 shares compared to other sizes. Verousis and ap Gwilym (2013) find trade size clusters at multiples of 500 shares on the London Stock Exchange. Mahmoodzadeh and Gençay (2017) document the human's preference for round prices after exchanges change their decimal price systems. Clustering is also identified and analyzed in foreign exchanges (Moulton, 2005), derivative markets (ap Gwilym & Meng, 2010) and the U.S. equity market (Ikenberry & Weston, 2008).

The reason for clustering can be explained by human's preference for round numbers. Individuals tend to use rounded numbers as cognitive reference points (Rosch, 1975) to simplify and save effort in the decision-making and evaluation process (Ikenberry & Weston, 2008; Kuo et al., 2015; Lacetera, Pope, & Sydnor, 2012). Therefore, the cognitive reference of round numbers sets human trades apart from robot trades (Mahmoodzadeh & Gençay, 2017; O hara, Yao, & Ye, 2014). Moreover, rounded numbers are more comfortable to recall and input. Therefore, orders with round numbers have a higher probability of being placed and executed, which leads to the phenomena of trades cluster at rounded numbers.

As increasing public attention has drawn more and more retail investors to participate in cryptocurrency trading worldwide (Bogart, 2019; FCA, 2019; Henry et al., 2019), we also expect to observe behavioral regularity in cryptocurrency markets. Wash trades, however, would distort this characteristic for the following reason. For the efficiency and quantity of fake trades, fraudulent crypto exchanges are likely to use machine-based automated trading

programs for wash trading due to savings on manpower compensation and its good efficiency, especially when fake orders feature small trade-size but large quantities for concealment. The psychological and behavioral regularity of humans is barely the priority to factor into the constant manipulation of trade history systematically, and therefore not reflected in the fake trading volume. Consequently, the wash trading inflates machine-traded volume and naturally reduces the proportion of human-traded volume. Therefore, the behavior regularity should be distinguished between wash trading and non-wash trading exchanges.

Because most cryptocurrencies can be traded in fractions, and some currencies have high unit value (especially BTC), in the following section, we set the smallest unit, the base unit, as one unit in a certain decimal place valued in the neighborhood of one US dollar. For instance, with the price of Bitcoin varying around \$8000-\$10000, most BTC-USD orders are below 1 BTC, therefore, round numbers in traditional financial markets such as 100, 1000 or 10000 are too large for individual traders. The value of 10^{-4} BTC is in the order of magnitude of one US Dollar, therefore, it is natural to consider 10^{-4} BTC as the base unit in this study. Similarly, the base unit of ETH, LTC and XRP are 0.0001ETH, 0.01LTC and 1XRP, respectively. We then examine whether trade-size clustering appears at multiples of 100 base units for each cryptocurrency.

4.2.1 Histograms of Trade Size

In this section, we identify the trade-size clustering by analyzing the probability distribution of trade size across exchanges. Figure 2 depicts trade size distributions of representative exchanges in two observation ranges for BTC, ETH, LTC and XRP, highlighting the clustering effect at the rounded trade sizes.[22] Histograms of remaining exchanges are shown in Appendix B. Panel R, Panel UT and Panel U depict the trade-size distribution for regulated exchanges, unregulated Tier-1 exchanges, and unregulated Tier-2 exchanges, respectively. Note that the Y-axis represents the probability that transactions fall into each interval, shown on a log scale.

Firstly, three regulated exchanges display a downward sloping curve with prominent peaks at multiples of 5000 base units in the range of 0-10 BTC (e.g., 0.5 BTC, 1 BTC, 1.5BTC, 2BTC, etc.). Similar patterns also appear in distributions of ETH, LTC and XRP. The findings suggest the presence of trade size clustering in regulated crypto-exchanges. This finding is consistent with the trade pattern in regulated financial markets, which display a

[22] The observation ranges include 0-1BTC, 0-10BTC, 0-10ETH, 0-100ETH, 0-100LTC, 0-1000LTC, 0-10,000XPR and 0-100,000XPR.

downward trend[23] with a trade size clustering effect, see for example (Alexander & Peterson, 2007; ap Gwilym & Meng, 2010; Mahmoodzadeh & Gençay, 2017; Verousis & ap Gwilym, 2013). Similar to participants in traditional markets, cryptocurrency investors show their preference in rounded trade size.

In unregulated Tier-1 exchanges, clustering patterns are more noticeable in the trade-size distribution with the small observation window. Take Bitcoin as an example, UT6 and UT10 in Figure 2 are two Tier-1 exchanges that do not show clear clustering patterns. Besides, most trades of UT6 are concentrated at smaller sizes and display an indefinable drop in frequency, especially in LTC and XRP trades. Moreover, clustering patterns in different assets vary across crypto exchanges and have shown no overall connections. For some Tier-1 exchanges, clustering is less apparent in the trades of XRP than other cryptocurrency trading (see Panel UT2, UT4 and UT5 of Appendix B).

[Insert Figure 2]

In unregulated Tier-2 exchanges, we barely observe any apparent clustering peaks at round sizes in any cryptocurrencies. Moreover, trade patterns vary dramatically and are distinguishable from the typical downward distribution. Some exchanges show occasional peaks at specific round numbers. For instance, in U5, trades with 0.1 BTC are predominant in the distribution. Additionally, trade frequency does not monotonically change with the increase in trade size in all cryptocurrency trades of U5. Similar issue is observed in U7, U8, U9, U13 and U16 (see Appendix B).[24] In U11, BTC trades tend to cluster at multiples of 1000 units (e.g., 0.1 BTC, 0.2 BTC, etc.) in the range of 0-1BTC but not others. The trade patterns of U11 display inexplicable bulges for ETH, LTC and XRP. Visually, U14 shows scarcely conspicuous peaks at the rounded size of all cryptocurrency trades. A uniform distribution is observed in LTC and XRP, as well as large observation range of BTC and ETH. The finding indicates that investors trade with nearly equal frequency at different trade sizes, which is against the intuition and regularity in financial markets. Furthermore, at least six Tier-2 exchanges display uniform patterns in cryptocurrency trades.[25]

Therefore, trades in unregulated Tier-2 exchanges do not reflect much of humans' natural tendency towards clustering in trade sizes, therefore implying suspicious trading behavior in these exchanges. Moreover, these erratic patterns also reveal unpredictable and likely manipulated trading features and behaviors in unregulated Tier-2 exchanges.

[23] Downward trade distribution is attributable to the fact that large-size orders are less frequently placed and executed.
[24] See examples in Appendix B: U7, U9, and U15 in BTC trades; U3, U7, U8, U9 and U15 in ETH trades; U9 in XRP trades
[25] See example U1, U2, U4,U6, U10,U12 in Appendix B

4.2.2 Statistical Tests for Clustering

To quantitatively investigate the effect of trade-size clustering, we conduct the Student's *t*-test for each crypto exchange and compare the trade frequencies at rounded trade size with the highest frequency of nearby unrounded trades. For each trading pair, we set up two sets of observation windows: windows centered on multiples of 100 units (100X) with a radius of 50 units (100X-50, 100X+50), and windows centered on multiples of 500 units (500Y) with a radius of 100 units (500Y-100, 500Y+100).[26] Trade frequency is calculated as the number of trades with size i over total trade numbers in the observation window. For example (see Figure 3), in the BTC trades of R1, the observation window around 200 units (0.02BTC) ranges from 150 units (0.015BTC) to 250 units (0.025 BTC). Trades at 0.02BTC constitute 16.42% of total trades in 0.015-0.025BTC while the highest trade frequency of unrounded trades is only 2.54% in the observation range. The apparent difference indicates that trades with 0.015-0.025BTC clustering at 0.02BTC (200 base units).

[Insert Figure 3]

[Insert Table 3]

Table 3 presents the *t-test* results for size clustering in regulated exchanges (Panel A), unregulated Tier-1 (Panel B) and Tier-2 exchanges (Panel C). As expected, in all three regulated exchanges (Panel A in Table 3), trade frequency at round size is strikingly higher than non-round ones regardless of cryptocurrencies and observation ranges, which is consistent with our findings in Figure 2. Additionally, size clustering is observed more apparently at multiple of 500 units in terms of difference and t-statistics. For example, in BTC trades of exchange R1, the difference in frequency is 9.09% in trade size increments of 100 units (e.g. 0.01 BTC, 0.02 BTC, and 0.03 BTC) while the difference is 20.28% at the size which is the common multiple of 500 units (e.g. 0.05BTC, 0.01 BTC, 0.015 BTC). The results are consistent with the human bias of round numbers especially multiples of five.

Similar to regulated exchanges, three unregulated Tier-1 exchanges (UT3, UT7 and UT9) show positive and significant differences in trades of all available cryptocurrencies, implying a clustering effect.[27] Moreover, trade clustering appears more frequently at multiples of 500 units: for example, six Tier-1 exchanges (UT1, UT3, UT5, UT7, UT8 and UT9) exhibit noticeable clustering effects at multiples of 500 units for all four cryptocurrencies. However, UT6 and UT10 show insignificant differences of frequencies between rounded and unrounded sizes, indicating a likelihood of washing trades.

[26] Because orders in multiple exchanges have different decimal place accuracy, we round trade history data to the base unit for consistency through all exchanges.
[27] All differences are statistically significant at 1% level, except XRP trades of UT9 which is at 5%.

Contrast to unregulated Tier-1 exchanges, the clustering effect is less observed in unregulated Tier-2 exchanges. Half of exchanges exhibit no sign of clustering for all cryptocurrencies in both observation windows (100X; 500X). Except U13, all Tier-2 exchanges have no clustering in at least one cryptocurrency. In BTC and ETH trades, apparent size-clustering is observed in more Tier-2 exchanges (six exchanges) than number of exchanges in trades of LTC (four exchanges) and XRP (one exchange). Besides, in some exchanges, trade cluster becomes less apparent at higher level of roundness (multiple of 500 units). For example, in U3 and U5, frequencies at multiples of 100 units are higher, significantly at 1% level but clusters at multiples of 500 units are not significant.

We also regress the (logit) percentage of trades at a certain size on various dummy variables which are set to one at rounded sizes. Moreover, we apply Chi-squared tests to compare whether the trade distributions of unregulated exchanges are in line with patterns in regulated exchanges. The results (shown in Appendix C and Appendix D) are consistent with the tests in this section.

In sum, in this section, we document that regulated exchanges display an evident clustering effect in trade size, whereas 30% and 50% of unregulated Tier-1 and Tier-2 exchanges, respectively, display no trade-size clustering in all cryptocurrencies.

4.3 Tail distribution

In this section, we examine the tails of trade-size distributions for each crypto exchange. By fitting the tail parts with power-law distribution, we investigate the consistency with the empirical regularity of financial markets. Inconsistency with the power law suggests the existence of manipulative behavior in cryptocurrency trades.

4.3.1 Power-law distribution as a statistical and behavioral benchmark

Power law features a wide variety of physical, biological, and man-made phenomena approximately. In economics and finance, power-law distribution appears to be a common property of regularities with 'fat tail' property. For example, Pareto distribution of income is a particular case of power law (Pareto, 1896). In equity markets, power-law tails are observed in the distribution of stock return (Gopikrishnan et al., 1999), trade size (Gopikrishnan et al., 2000) and share volume (Plerou et al., 2000). Fluctuation in foreign exchange market also follows a power-law distribution (Da Silva, Matsushita, Gleria, & Figueiredo, 2007; Ohnishi et al., 2008; Vandewalle, Ausloos, & Boveroux, 1997). In cryptocurrency market, M.-Y. Li et al. (2019) and Schnaubelt et al. (2019) document that

the trade-size tail distribution for bitcoin follows a power law with exponents of 1.75 and 1.44, respectively.

One explanation for power-law tails in empirical regularities is the trading behavior of large investors, who try to avoid large impact in price movement (Gabaix, Gopikrishnan, Plerou, & Stanley, 2003a). Other studies attribute the emergence of power-law to the investors' limited information on value of assets (Kostanjčar & Jeren, 2013; Nirei, Stachurski, & Watanabe, 2018) and herding behavior of traders (Nirei et al., 2018).

Statistically, power-law distribution has cumulative density function (CDF) that follows the form

$$\mathrm{P(X > x)} \sim \mathrm{x}^{-\alpha} \tag{2}$$

where α is known as the power-law exponent or tail exponent and estimator for probability density function (PDF) is $\alpha + 1$.

As mentioned before, the participation of large investors and asymmetry information determines traders' behavior, resulting in power-law distribution. In the crypto market, large participants (e.g. institutional investors or large retail investors) have increasingly engaged in the cryptocurrency trade over the past few years. And retail traders have asymmetric information on the value of cryptocurrency. Therefore, fluctuation in cryptocurrency trades is highly likely to conform to power law, which has been proven true for bitcoin, shown in Schnaubelt et al. (2019) and M.-Y. Li et al. (2019). On the other hand, manipulated records exhibit different patterns and may not display this statistical property in the tails.

4.3.2 Power law and tail exponents

To examine trade size distribution tails, we used two widely adopted techniques: The first one is to take the logarithm of the empirical probability density function and fit the log-log data to power-law distribution by Ordinary Least Square (OLS). The second one is to apply the Maximum Likelihood Estimation approach (MLE) and use Hill estimator $\hat{\alpha}_{Hill}$ for the data fitting. Hill estimator is asymptotically normal and calculated as follows: (Clauset, Shalizi, & Newman, 2009; B. M. Hill, 1975)

$$\hat{\alpha}_{Hill} = n\left(\sum_{i=1}^{n} ln\frac{x_i}{x_{min}}\right) \tag{3}$$

where n is the number of observation and x_{min} is the cut-off threshold. The distribution yield to power-law after x_{min}. In this study, trade size distributions are constructed for

empirical probability density function. The cut-off x_{min}, which signifies the start of tails, is set as the top 10% largest trades during sampling period.

Then the fitting results are checked for whether the values of exponent α suit the Pareto–Lévy range ($1 < \alpha \leq 2$). Gabaix, Gopikrishnan, Plerou, & Stanley (2003b) show that stock trade size follows a half cubic law ($\alpha = 1.5$) both theoretically and empirically. Various studies in financial-asset trading volumes or sizes have shown that the vast majority of tail exponents lies in the Pareto–Lévy regime ($1 < \alpha < 2$) for traditional financial assets[28] and for bitcoins (M.-Y. Li et al., 2019; Schnaubelt et al., 2019).

Table 4 presents the results from OLS and MLE fittings for four cryptocurrency trades of fitted exchanges. We then visually inspect the goodness of fit and identify whether crypto exchange display a power-law tail in trade size distribution, shown in Figure 4.

[Insert Table 4]

As expected, in regulated exchanges, both scaling estimators $\hat{\alpha}_{OLS}$ and $\hat{\alpha}_{Hill}$ lie in the Pareto–Lévy regime and suggests stable a power-law decay in all cryptocurrency trades. Similar patterns are observed in 50% of unregulated Tier-1 exchanges. In contrast, estimators of two Tier-1 exchanges (UT4 and UT5) do not fall into the Pareto–Lévy range for four cryptocurrencies and suggest inconsistency with power-law exponents for trade size in traditional markets. Besides, tail exponents of UT7, UT8 and UT10 are outside the range of 1 to 2 in one cryptocurrency.

In unregulated Tier-2 exchanges, only three exchanges show estimated exponents within Pareto–Lévy range, whereas 62.5% show statistical evidence in disconformity to parameters of empirical regularity in four cryptocurrencies. The rest, the estimated exponents of U12 follow Pareto–Lévy range in LTC and ETH trades while U14 and U16 show similar fashion in LTC and ETH trades, respectively.

[Insert Figure 4]

Figure 4 displays the probability density for trade size and the fitted power-law distributions on log-log plots, with one regulated and six unregulated exchanges as representatives for brevity. Figures of the entire sample are listed in Appendix E.

[28] Gopikrishnan et al. (2000) find that the power law exponent of trade volume is around 1.5 in US equity market. Plerou and Stanley (2007) investigate trades in New York Stock Exchange, London Stock Exchange and Paris Bourse and show that trade size in all three markets display power law decay with exponent in the range from 1 to 2. Moreover, value of exponents is not affected by industry and market capitalization. Note that Mandelbrot (1960) propose that income follows the stable "Pareto–Lévy" distributions with $1 < \alpha < 2$.

Similar to mainstream financial markets, transactions from regulated exchanges display a downward linear trend in the log-log plots and appear visually to fit the power-law distribution. For instance, in Panel R1 of Figure 4, empirical data points fall around the fitted lines without obvious outliners, implying that trades in regulated exchange generally follow the power law in all four listed cryptocurrencies. Generally speaking, the OLS line fits equally in the whole range while MLE estimation weighs more at the start of the tail, where the probability value is higher. Consistent with regulated exchanges, 90% of unregulated Tier-1 exchanges resemble power-law tails in trade size distributions. Straight lines estimated by OLS and MLE are roughly fitted to the data. Conversely, UT6 (shown in Figure 4) shows curvy shape in tails and fails to show power-law distribution in the trade size.

In unregulated Tier-2 exchanges, tail distributions vary differently and display irregular patterns across exchanges and cryptocurrencies. Five Tier-2 exchanges (U6; U9; U13; U15; U16) show a linear decrease in the tail zones and comply with the power-law tail. In U3 (shown in Figure 4), data points are step-like distributed on the logarithm scale in trades of BTC, ETH and LTC while tail in XRP trade distribution is horizontal at the start, followed by a slower decay than power law predicted. U4 displays a concave curve in all cryptocurrency trades. We also observe similar tail behaviors in BTC, LTC and XRP trades of U10. In U12, data points are not randomly distributed around the fitted line, either clustered above or below the line. In BTC trades of U14, the tail appears to be level with some outliers far from the line. ETH, LTC and XRP trades of U14 show a step-like decay.

Combined results above, regulated exchanges decay as power law predicts, and show estimators consistent with Pareto–Lévy exponents in financial markets. 50% of Tier-1 exchanges display power-law tail with exponents characterized by the Pareto–Lévy regime in all cryptocurrencies. 75% of unregulated Tier-2 exchanges fail to follow the Pareto–Lévy power law in trade size distribution as regularities in financial markets, implying that the presence of suspicious wash trading.

4.4 Quantifying Wash Trading

Sections 4.1 to 4.3 suggest the absence of suspicious trades in regulated crypto exchanges, however, present idiosyncratic and discrepant patterns in unregulated exchanges (especially Tier-2). In this section, we develop a way to estimate the amount of wash trading for unregulated exchanges by taking the regulated exchanges as a benchmark. Particularly, humans tend to place orders with round sizes, in contrast, unrounded trades normally relate to program trading of various purposes, such as market marking, high-frequency arbitration and more importantly wash trading, which is highly likely conducted using automated programs considering the efficiency and quantity of trade orders required. Therefore,

rounded trades/unrounded trades can be used as a reasonable proxy for human orders/computer trades. The main idea of wash-trade estimation is first to find the benchmark ratio of unrounded to rounded trades in the regulated exchanges; assuming unregulated exchanges follow the same benchmark ratio of legitimate (non-wash) unrounded trades to rounded trades, we can estimate the legitimate amount of unrounded trades from rounded trades for unregulated exchanges. The difference between the observed unrounded trading volume and "legitimate" trading volume is regarded as the wash-trading volume. Note that since it is nearly impossible to directly label wash trades without confession of exchanges, our method provides a rough but feasible way of estimating wash trades.

4.4.1 Trade-size Roundness

First, we test level of roundness to compare trade sizes of unregulated exchanges with regulated ones. The level of roundness is a qualitative parameter describing the decimal or integer places of last non-zero digit. For instance, 1.01BTCs have a higher level of roundness than 2.123BTC; 100ETHs have a higher level of roundness than 1234ETH.[29]

Retail trade sizes should display a higher level of roundness than the artificial ones by computer programs. We thus expect regulated exchanges to present a higher level of roundness in trade sizes compared with exchanges doing wash trades. For each crypto-exchange, we first analyze the trade-size distribution over levels of roundness (ten thousands, thousands, hundreds, tens, ones, tenths, hundredths etc.) in the base unit framework. We then apply the average frequency distribution of regulated exchanges as a benchmark. The distribution for the level of roundness is then compared between regulated and unregulated exchanges.

[Insert Table 5]

The average distribution of regulated exchanges is used as a benchmark and compared with unregulated exchanges. Table 5 shows the Chi-squared test for comparison between regulated and unregulated exchanges for four cryptocurrencies. All Tier-1 exchanges have significantly large Chi-squared statistics in at least one cryptocurrency. As of unregulated Tier-2 exchanges, except U7 in BTC trades, all trades show completely different roundness distributions from regulated exchanges with a 1% significance level nearly for all cryptocurrencies.[30] This finding shows that unregulated exchanges, especially unregulated

[29] For 1.01BTC, the place value of last non-zero digit (1) is hundredths, while the place value of last non-zero digit (3) is thousandths in 2.123 BTC. In 100 ETH, the place value of last non-zero digit (1) is hundreds while the place value of last non-zero digit (4) is ones.

[30] The different of U9 in BTC trades, U11 and U16 in ETH trades are significant at 5% level.

Tier-2 exchanges, present a lower level of roundness in trade size relative to the regulated exchanges.

4.4.2 Volume of Wash Trades

We estimate the number of wash trades by calculating the abnormal proportion of unrounded trades for various exchanges. Specifically, we categorize trading volumes into round and non-round ones by checking if the last non-zero digit of a certain trade size is less than 100 basis units or not. In regulated crypto exchanges, rounded trades constitute around 30% of total trades, which complies with volume of human traders in US equity markets (Gomber, Gsell, Pujol, & Wranik, 2009; Tabb, Iati, & Sussman, 2009). Therefore, the classification of rounded and unrounded trades could be used to differentiate human and automated trading for cryptocurrencies. We assume that regulated exchanges do not have wash trades and that the computer-based legitimate (non-wash) trades in unregulated exchanges have the same (log) ratio to the human trades as those in regulated exchanges, we, therefore, estimate the wash trading volume in unregulated exchanges as the total unrounded volume deducted by legitimate (non-wash) trades estimated from rounded trades.

We perform a pooled regression to estimate the ratio of (log) unrounded volume to (log) rounded volume for all regulated exchanges with a weekly frequency:

$$\ln\left(\mathrm{V}_{Unround\,jt}\right) = \alpha + \beta * \ln\left(\mathrm{V}_{Round\,jt}\right) + \mu_{jt} \tag{4}$$

where $\mathrm{V}_{Unround\,jt}$ and $\mathrm{V}_{Round\,jt}$ are unrounded and rounded trading volumes of regulated exchange j at week t. We employ the parameters in (4) to calculate the legitimate (non-wash) unrounded trades of unregulated exchanges using their corresponding rounded trading volumes. Wash trade volumes are thus calculated as the difference between total unrounded trades and legitimate unrounded trades.

[Insert Table 6]

Table 6 presents the average percentage of wash trades, weighted by volume of four cryptocurrency pairs, for all exchanges. On average, wash trades account for an average of 70% of total trading volume in all unregulated exchanges, around 53% for Tier-1 and 80% for Tier-2 exchanges. Our estimates are consistent with the previous findings that wash trading tends to appear more in less-regulated exchanges.

4.5 Conclusive Evidence of Wash Trading

We summarize all the statistical results (see Appendix F and G) and report failure/pass of tests in each cryptocurrency for crypto exchanges in Appendix F, including the Chi-squared test for Benford's Law, *t*-test for trade-size clustering, and tail exponents for the power law. The rest of tests are summarized in Appendix G, in which we compare the trade size distribution and roundness level of unregulated exchanges with patterns of regulated exchanges. Figure Figure illustrates the percentage of failed tests by exchanges and types of cryptocurrencies. For each exchange, percentage of failed tests is measured as the number of failed tests at a 5% significance level over total number of tests of all four trading assets. Similarly, percentage of failed tests by cryptocurrency is calculated as the number of failed tests at a 5% significance level over total number of tests in each type of cryptocurrency.

[Insert Figure *]*

Overall, regulated exchanges (R1, R2, and R3) have passed all sets of tests, likely due to the heavy regulation from the New York State. Unregulated Tier-1 exchange UT9 and Tier-2 exchange U13 also show consistency with empirical regularities in tests for first leading digits, clustering and tail distribution. However, compared with regulated exchanges, UT9 and U13 display different patterns in trade size distribution and level of roundness (Appendix G).

In general, unregulated Tier-1 exchanges have lower failure rates than unregulated Tier-2 exchanges. Some Tier-1 exchanges show mildly skeptical patterns in wash trading. It is reasonable, as being well-established businesses in the industry, wash trading accusations have the potential to damage their business goodwill seriously. Some of the unregulated Tier-1 firms might have already been doing compliance work in other jurisdictions than the U.S.

However, except for one Tier-2 exchange (U13), all Tier-2 exchanges fail for at least one-third of statistical tests, indicating that trade patterns of Tier-2 exchanges are inconsistent with regularity in regulated exchange as well as traditional financial markets. Therefore, those exchanges are highly like to engage in wash trading.

[Insert Table 7]

Grouped by cryptocurrency, the percentage of failed tests is highest in XRP trades (54.17%), followed by BTC (47.44%), LTC (46.97%), and ETH (42.31%). In terms of daily trading volumes, XRP is significantly less than other cryptocurrencies, implying that manipulation is more likely to occur in less traded markets since inactive markets are less monitored and manipulation is easier to carry out.

We also examine the relationship between failed rates and fractions of wash trades as in Table . Percentage of wash trade is positively associated with percentage of failure at 1% significance level, 1% increase in the failure rates corresponds to a 0.597% more wash trading percentage. This suggests that our indicator for wash trade does reflect the suspicious trading activities in unregulated exchanges.

5 Wash Trading Incentives, Exchange Ranking, and Regulatory Implications

5.1 Wash Trades and Exchange Ranks

Retail investors rely on third-party rating or ranking websites to decide which crypto exchange to use. Data providers or ranking agencies, especially those attracting a large number of web traffic, therefore play an important role in exchanges' customer acquisition.

We acquire proprietary data from Coinmarketcap.com. The platform started business on providing crypto market capitalizations, pricing, and other information on all kinds of crypto currencies. Growing together with the industry, the company has developed into a top data provider and ranking agency in the industry. By 2020, it serves 4.2million unique visitors around the globe for 32.6 million visits per month, dominating its kind.[31] It is so success that Binance proposed an acquisition in March 2020. The deal is undisclosed but rumour says $400 Million. [32] Currently, the 'Crypto Standard & Poor s' declared itself as accurate and neutral. Despite some accusations on its ranking system every now and then, there is no evidence of crypto rating services attempting to manipulate price. However, the exchange ranking system and its influence over millions of users is not to be ignored.

Currently, there is little regulation on data providers and rating agencies in the cryptocurrency industry around the world. Given their influence and vital function, these third party rating agencies are likely to face more regulation just like credit rating agencies in traditional financial markets, , there is no doubt proper regulation is required like credit rating agencies in traditional finance market.

To study the incentives of exchange wash trading, we quantitatively assess the relationship between trading volume and exchange rank. In Figure 6, we observe a noticeably downward linear trend with very large goodness of fit (adjusted R^2 of 93%), indicating that trading volume is a major factor in ranking exchanges. An increase in trade volume boosts exchange ranking, specifically, a 10% increase of trading volume can move exchange rank up by 2

[31] Data acquired from Similiarweb, accessed on Jun/12/2020. https://www.similarweb.com/
[32] Bitcoin And Crypto World Rocked By An Estimated $400 Million Binance Bid For CoinMarketCap—Report, 2020.Forbes. https://www.forbes.com/sites/billybambrough/2020/03/31/bitcoin-and-crypto-world-rocked-by-massive-400-million-binance-bid-for-coinmarketcap-report/

places. Based on results from previous sections, unregulated exchange has wash trading volume that constitute 70% of total volume on average. According to Figure 6, this moves the rank up about 46 places.

This finding supports the intuition that the main motivation of wash trading is to gain prominence and market share so that the exchange can generate higher profits. By wash trading, exchanges inflate trading volume and therefore rank higher, resulting in wider public awareness, larger user base, and finally higher revenue on transaction fees. Additional revenue comes from margin interest, advertisement and consultant fees (also known as listing fees, where exchanges charge developers to list their cryptocurrency). Overall, the profit crucially depends on brand awareness and website traffic as two critical factors for customer acquisition, both of which heavily rely on public rankings in broadly recognized data tracking/ranking services or third-party websites such as CoinMarketCap. Given that the ranking algorithm is based on reported trading volume, many crypto exchanges are naturally suspected of undertaking illicit activities in order to attract users and to survive in face of fierce competition. .

[*Insert* Figure]

5.2 Implications for Regulation and Industry Practice

Wash trading is legally prohibited in most developed financial markets. For example United States has banned wash trading in the Commodity Exchange Act (CEA) 1936, European Union listed it in the Market Abuse Directive No 2003/6/EC etc. Therefore financial services that are operating under traditional regulation framework are naturally prohibited from wash trading. Laws and regulation specific to cryptocurrencies are still being crafted around the globe; after all, the majority of crypto exchanges are just a few years old.

Against such backdrop, crypto wash trading has become a major concern. Several companies in the industry started release industry reports and expose wash trading of crypto exchanges. However, these reports tend to rely on anecdotes and cases. For instance, the Bitwise report was generated as a descriptive presentation for SEC regulators (Fusaro & Hougan, 2019). The number of exchanges observed is limited and most of the data come from a short time series (between Mar 4 and Mar 8 in 2019). Another related example is the report by Alameda (Alameda, 2019), a quantitative trading firm. The methodologies applied involve manually checking the website of exchange webpages, and using their internal trading data to compare results, not to mention that Alameda is closely connected with a crypto exchange itself, which casts some doubts on the impartiality of the report.

Besides creating reports, some industry leaders took action to fight wash trading problem. CoinMarketCap, for example, introduced a mandatory API program for all listed exchanges

to improve credibility and transparency.[33] They later developed another set of rank algorithm based on exchanges liquidity instead of volume.[34] CryptoCompare, a British cryptocurrency data analysis firm, launched a unique exchange benchmark product which would help safeguard against false exchange volume reports.[35] Nomics, a data provider developed Transparency Volume based on their ranking criteria, claiming it is less likely to include wash trading and other misleading forms of volume.[36] To our knowledge, we provide the first academic study to systematically and rigorously analyze crypto wash trading as an industry-wide phenomenon. The statistical tests add to the toolkit for detecting wash trading and regulation.

Importantly, we show that regulated and unregulated exchanges exhibit vastly divergent trading behaviors. Regulated exchanges pass all tests and the trading history matches theories oand patterns in traditional financial markets that are relatively free from wash trading. In contrast, unregulated Tier 1 exchanges on average failed 26% of the tests, which shows signs of self-regulation and reputation maintenance. Even more glaringly, unregulated Tier 2 exchanges failed 65% of all tests on average, which means highly suspicious training history.

Our findings imply that regulation makes a huge impact in cryptocurrency markets, with ramifications on investor protection, price discovery, and financial stability. Moreover, it offers a set of tools to convincingly unveil wash trading of exchanges, which can be applied for combating non-compliant and unethical behaves.

6 Conclusion

The decentralized nature of blockchains and the nascent status of cryptocurrencies provide a special environment where we observe both regulated and unregulated exchanges that wield tremendous power and influence in the ecosystem. In this paper, we show that unregulated crypto exchanges are highly likely involved in wash trading. To this end, we examine the three aspects of trade sizes, i.e. leading digits, clustering pattern (including roundness) and tail distribution for different exchanges, and compare these aspects between regulated exchanges (with Bitlicense and under the preview of New York State Department of Financial Services) and unregulated exchanges.

First-digit distributions of trade size follow the Benford's law for regulated exchanges, whereas nearly 30% of unregulated exchanges show violations. Furthermore, regulated exchanges show apparent trade clustering at round sizes and a high level of transaction roundness. However, the levels of roundness are generally low and the trade-size clustering phenomenon is less prominent for unregulated exchanges. Finally, regulated exchanges display power law decay with tail exponents characterized by Pareto–Lévy range, consistent

[33] https://blog.coinmarketcap.com/2019/05/01/happy-6th-birthday-data-alliance-block-explorers-and-more/
[34] https://blog.coinmarketcap.com/2019/11/11/coinmarketcap-launches-new-liquidity-metric/
[35] https://blog.cryptocompare.com/introducing-cryptocompares-exchange-benchmarking-methodology-f44adc506431?gi=c675f5ca6bef
[36] https://blog.nomics.com/essays/transparent-volume/

with regularity in financial markets. Conversely, 20% of Tier-1 and 75 % of Tier-2 exchanges fail to follow Pareto–Lévy law in trade-size distribution of any cryptocurrency. We also perform a rough estimation on the wash trades fraction: 52.5% of trade volume is attributed to wash trading in unregulated Tier-1 exchanges while 80.5% in Tier-2 exchanges. We further show that wash trading inflates rankings from widely recognized platforms and therefore distorts investors' attention.

Overall, our findings provide support for the view that a large proportion of unregulated exchanges inflate trade volume through computer programs. Wash trading by crypto exchanges therefore distorts demand and supply of cryptocurrencies and artificially affects the price and harms the development of the cryptocurrency market in the long run. Given that wash trade seems prevalent in unregulated crypto-exchanges, in which newcomers and less reputable exchanges engage in more suspicious activities, our findings provide a cautionary tale to regulators around the globe. At the same time, we remind the readers of the tremendous effect of regulation and provide tools to strengthen the regulations for crypto exchanges.

References

Aggarwal, R. K., & Wu, G. (2006). Stock market manipulations. *The Journal of Business, 79*(4), 1915-1953.

Alameda. (2019). *Investigation into the Legitimacy of Reported Cryptocurrency Exchange Volume*. Retrieved from

Alexander, G. J., & Peterson, M. A. (2007). An analysis of trade-size clustering and its relation to stealth trading. *Journal of Financial Economics, 84*(2), 435-471.

Alexandre, A. (2019). Polish Exchange Shuts Down and Disappears With Customers Funds. Retrieved from https://cointelegraph.com/news/report-polish-exchange-shuts-down-and-disappears-with-customers-funds

Allen, F., & Gale, D. (1992). Stock-price manipulation. *The Review of Financial Studies, 5*(3), 503-529.

ap Gwilym, O., & Meng, L. (2010). Size clustering in the FTSE100 index futures market. *Journal of Futures Markets: Futures, Options, and Other Derivative Products, 30*(5), 432-443.

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American philosophical society*, 551-572.

Blandin, A., Cloots, A. S., Hussain, H., Rauchs, M., Saleuddin, R., Allen, J. G., . . . Zhang, B. Z. (2019). Global cryptoasset regulatory landscape study. *Available at SSRN*.

Bogart, S. (2019). Blockchain Capital Bitcoin Survey. Retrieved from https://medium.com/blockchain-capital-blog/bitcoin-is-a-demographic-mega-trend-data-analysis-160d2f7731e5?

Brett, J. (2019). Congress Considers Federal Crypto Regulators In New Cryptocurrency Act Of 2020. Retrieved from https://www.forbes.com/sites/jasonbrett/2019/12/19/congress-considers-federal-crypto-regulators-in-new-cryptocurrency-act-of-2020/

Chen, T. (2018). Round-number biases and informed trading in global markets. *Journal of Business Research, 92*, 105-117.

Christie, W. G., & Schultz, P. H. (1994). Why do NASDAQ market makers avoid odd‐eighth quotes? *The Journal of Finance, 49*(5), 1813-1840.

Clarkson, P., Nekrasov, A., Simon, A., & Tutticci, I. (2015). Target price forecasts: Fundamental and non-fundamental factors. *Available at SSRN 2104433*.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review, 51*(4), 661-703.

Cong, L. W., Li, Y., & Wang, N. (2019a). Tokenomics and platform finance. *Available at SSRN 3472481*.

Cong, L. W., Li, Y., & Wang, N. (2019b). Tokenomics: Dynamic adoption and valuation. *Review of Financial Studies, Forthcoming*.

Cong, L. W., Li, Y., & Xiao, Y. (2020). A Brief Introduction to Tokenomics. Book Chapter in *Palgrave Handbook of FinTech and Blockchain,* (edited by Maurizio Pompella and Roman Matousek, Palgrave MacMillan, *Forthcoming*.

Cong, L. W., He, Z., & Li, J. (2020). Decentralized Mining in Centralized Pools. *Review of Financial Studies, Forthcoming.*

Cong, L. W., Miao, Q., Tang, K., & Xie, D. (2019). Survival Scale: Marketplace Lending and Asymmetric Network Effects. *Available at SSRN 3461893*.

Cumming, D., Johan, S., & Li, D. (2011). Exchange trading rules and stock market liquidity. *Journal of Financial Economics, 99*(3), 651-671.

Da Silva, S., Matsushita, R., Gleria, I., & Figueiredo, A. (2007). Hurst exponents, power laws, and efficiency in the Brazilian foreign exchange market. *Economics Bulletin, 7*(1), 1-11.

Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting, 5*(1), 17-34.

FCA. (2019). Cryptoassets: Ownership and attitudes in the UK. (Financial Conduct Authority). Retrieved from https://www.fca.org.uk/publication/research/cryptoassets-ownership-attitudes-uk-consumer-survey-research-report.pdf

Fidelity. (2019). Institutional Investments in Digital Assets. Retrieved from https://s2.q4cdn.com/997146844/files/doc_news/archive/59439969-390c-4354-94a9-772219d0b8b9.pdf

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *The Review of Financial Studies, 32*(5), 1798-1853.

Fusaro, T., & Hougan, M. (2019). Presentation to the US Securities and Exchange Commission. *Bitwise Asset Management*.

Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003a). A theory of power-law distributions in financial market fluctuations. *Nature, 423*(6937), 267-270.
Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003b). Understanding the cubic and half-cubic laws of financial fluctuations. *Physica A: Statistical Mechanics and its Applications, 324*(1-2), 1-5.
Gandal, N., Hamrick, J., Moore, T., & Oberman, T. (2018). Price manipulation in the Bitcoin ecosystem. *Journal of Monetary Economics, 95*, 86-96.
Gomber, P., Gsell, M., Pujol, G., & Wranik, A. (2009). Regulation and Technology in Equity Trading–The Impact of Regulatory and Technological Changes on Order Execution and the Trading Venue Landscape in Europe. *Facing New Regulatory Frameworks in Securities Trading in Europe*, 31-54.
Gonzalez-Garcia, J. (2009). *Benford s law and macroeconomic data quality*: International Monetary Fund.
Gopikrishnan, P., Plerou, V., Amaral, L. A. N., Meyer, M., & Stanley, H. E. (1999). Scaling of the distribution of fluctuations of financial market indices. *Physical Review E, 60*(5), 5305.
Gopikrishnan, P., Plerou, V., Gabaix, X., & Stanley, H. E. (2000). Statistical properties of share volume traded in financial markets. *Physical Review E, 62*(4), R4493.
Griffin, J. M., & Shams, A. (2018a). Is bitcoin really un-tethered? *Available at SSRN 3195066*.
Griffin, J. M., & Shams, A. (2018b). Manipulation in the VIX? *The Review of Financial Studies, 31*(4), 1377-1417.
Günnel, S., & Tödter, K.-H. (2009). Does Benford's Law hold in economic research and forecasting? *Empirica, 36*(3), 273-292.
Halaburda, H., & Gandal, N. (2016). Competition in the cryptocurrency market. *Available at SSRN 2506463*.
Henry, C. S., Huynh, K. P., & Nicholls, G. (2019). Bitcoin Awareness and Usage in Canada: An Update. *The Journal of Investing, 28*(3), 21-31.
Helms, K. (2020). $8.8 Trillion Traded in Cryptocurrency Spot and Futures Markets in Q1: Reports. Retrieved from https://news./trillion-traded-cryptocurrency-spot-futures-markets/
Hileman, G., & Rauchs, M. (2017). Global cryptocurrency benchmarking study. *Cambridge Centre for Alternative Finance, 33*.
Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 1163-1174.
Hill, T. P. (1995). Base-invariance implies Benford's law. *Proceedings of the American Mathematical Society, 123*(3), 887-895.
Hill, T. P. (1998). The first digit phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data. *American Scientist, 86*(4), 358-363.
Ikenberry, D. L., & Weston, J. P. (2008). Clustering in US stock prices after decimalisation. *European Financial Management, 14*(1), 30-54.
Imisiker, S., & Tas, B. K. O. (2018). Wash trades as a stock market manipulation tool. *Journal of behavioral and experimental finance, 20*, 92-98.
IOSCO. (2000). A Resolution on IASC Standards. In: Presidents Committee of IOSCO Madrid, Spain.
Jarrow, R. A. (1992). Market manipulation, bubbles, corners, and short squeezes. *Journal of financial and Quantitative Analysis, 27*(3), 311-336.
Kandel, S., Sarig, O., & Wohl, A. (2001). Do investors prefer round stock prices? Evidence from Israeli IPO auctions. *Journal of Banking & Finance, 25*(8), 1543-1551.
Kostanjčar, Z., & Jeren, B. (2013). Emergence of power-law and two-phase behavior in financial market fluctuations. *Advances in Complex Systems, 16*(01), 1350008.
Kuo, W.-Y., Lin, T.-C., & Zhao, J. (2015). Cognitive limitation and investment performance: Evidence from limit order clustering. *The Review of Financial Studies, 28*(3), 838-875.
Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review, 102*(5), 2206-2236.
Li, M.-Y., Cai, Q., Gu, G.-F., & Zhou, W.-X. (2019). Exponentially decayed double power-law distribution of Bitcoin trade sizes. *Physica A: Statistical Mechanics and its Applications, 535*, 122380.
Li, T., Shin, D., & Wang, B. (2019). Cryptocurrency pump-and-dump schemes. *Available at SSRN 3267041*.
Li, Z., Cong, L., & Wang, H. (2004). Discussion on Benford s Law and its Application. *arXiv preprint math/0408057*.

Mahmoodzadeh, S., & Gençay, R. (2017). Human vs. high-frequency traders, penny jumping, and tick size. *Journal of Banking & Finance, 85*, 69-82.
Mandelbrot, B. (1960). The Pareto-Levy law and the distribution of income. *International Economic Review, 1*(2), 79-106.
Mitchell, J. (2001). Clustering and psychological barriers: The importance of numbers. *Journal of Futures Markets: Futures, Options, and Other Derivative Products, 21*(5), 395-428.
Moore, T., & Christin, N. (2013). *Beware the middleman: Empirical analysis of Bitcoin-exchange risk.* Paper presented at the International Conference on Financial Cryptography and Data Security.
Moore, T., Christin, N., & Szurdi, J. (2018). Revisiting the risks of bitcoin currency exchange closure. *ACM Transactions on Internet Technology (TOIT), 18*(4), 1-18.
Moulton, P. C. (2005). You can't always get what you want: Trade-size clustering and quantity choice in liquidity. *Journal of Financial Economics, 78*(1), 89-119.
Nigrini, M. J. (1996). A taxpayer compliance application of Benford s law. *The Journal of the American Taxation Association, 18*(1), 72.
Nirei, M., Stachurski, J., & Watanabe, T. (2018). *Trade Clustering and Power Laws in Financial Markets*. Retrieved from
O hara, M., Yao, C., & Ye, M. (2014). What s not there: Odd lots and market data. *The Journal of Finance, 69*(5), 2199-2236.
Ohnishi, T., Takayasu, H., Ito, T., Hashimoto, Y., Watanabe, T., & Takayasu, M. (2008). Dynamics of quote and deal prices in the foreign exchange market. *Journal of Economic Interaction and Coordination, 3*(1), 99.
Palmer, D. (2020). Cobinhood Announces Shutdown, Claims It Will Audit User Accounts. Retrieved from https://www.coindesk.com/cobinhood-announces-shutdown-claiming-it-will-audit-user-accounts
Pareto, W. (1896). Course of political economy. In: Lausanne.
Pinkham, R. S. (1961). On the distribution of first significant digits. *The Annals of Mathematical Statistics, 32*(4), 1223-1230.
Plerou, V., Gopikrishnan, P., Amaral, L. A. N., Gabaix, X., & Stanley, H. E. (2000). Economic fluctuations and anomalous diffusion. *Physical Review E, 62*(3), R3023.
Plerou, V., & Stanley, H. E. (2007). Tests of scaling and universality of the distributions of trade size and share volume: Evidence from three distinct markets. *Physical Review E, 76*(4), 046109.
Ritter, J. R. (2008). Forensic finance. *Journal of Economic Perspectives, 22*(3), 127-147.
Roger, T., Roger, P., & Schatt, A. (2018). Behavioral bias in number processing: Evidence from analysts' expectations. *Journal of Economic Behavior & Organization, 149*, 315-331.
Rosch, E. (1975). Cognitive reference points. *Cognitive psychology, 7*(4), 532-547.
Sambridge, M., Tkalčić, H., & Jackson, A. (2010). Benford s law in the natural sciences. *Geophysical research letters, 37*(22).
Schnaubelt, M., Rende, J., & Krauss, C. (2019). Testing stylized facts of bitcoin limit order books. *Journal of Risk and Financial Management, 12*(1), 25.
Tabb, L., Iati, R., & Sussman, A. (2009). US equity high frequency trading: Strategies, sizing and market structure. *TABB Group report*(1), 2.
Vandewalle, N., Ausloos, M., & Boveroux, P. (1997). *Detrended fluctuation analysis of the foreign exchange market.* Paper presented at the Econophysic Workshop, Budapest, Hungary.
Verousis, T., & ap Gwilym, O. (2013). Trade size clustering and the cost of trading at the London Stock Exchange. *International Review of Financial Analysis, 27*, 91-102.
Yu, X. (2018). China to stamp out cryptocurrency trading completely with ban on foreign platforms. Retrieved from https://www.scmp.com/business/banking-finance/article/2132009/china-stamp-out-cryptocurrency-trading-completely-ban
Zhao, W. (2020). Crypto Exchange FCoin Insolvent After Revealing Up to $130M Bitcoin Shortfall. Retrieved from https://www.coindesk.com/crypto-exchange-fcoin-insolvent-after-revealing-up-to-130m-bitcoin-shortfall
Zitzewitz, E. (2012). Forensic economics. *Journal of Economic Literature, 50*(3), 731-769.

Table 1 Exchange information

Table 1 summarizes information of crypto exchanges in our sample. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank. Length of survival is the duration from establishment date to December 2019, categorized into three groups: "survived for five years and more", "length of survival between 2 and 5 years" and "survived for less than 2 years". Dollar trade volume is calculated as the sum of the product of trade size and contemporaneous price in dollars, include trading volume of BTC, ETH, LTC and XRP. SimilarWeb and Alexa are online platforms that provide rank and analysis according to website traffics. CoinMarketCap provides market capitalization and ranking of multiple cryptocurrencies and crypto exchanges.

			Ranking by Web Traffic			Ranking by Trade Volume
Exchange Code	**Age**	**Trade Volume ($mil)**	SimilarWeb Average Rank in Investment Section[37]	SimilarWeb Average Number of Monthly Visits[38] (millions)	Alexa[39]	CoinMarketCap[40]
Panel R Regulated exchanges						
R1	$\geqslant$ 5 year	10466	473	1.872	14297	63.7
R2	$\geqslant$ 5 year	15212	17	20.678	2254	50.3
R3	$\geqslant$ 5 year	1568	1418.5	0.487	23950	99.2
Panel UT Unregulated Tier-1 exchanges						
UT1	2year $\leqslant$ A<5year	41936	21	18.770	1630	10.5
UT2	$\geqslant$ 5 year	434	276	2.983	5960	89.9
UT3	$\geqslant$ 5 year	11175	345	2.757	9683	59.5
UT4	$\geqslant$ 5 year	34157	498.5	1.363	9815	27.9
UT5	$\geqslant$ 5 year	38789	285.5	1.673	8379	22.7
UT6	2year $\leqslant$ A<5year	4005	255.5	1.879	8663	55.2
UT7	$\geqslant$ 5 year	545	699	0.394	13357	53.3

[37] Ranking is based on report over the period from Aug 2019 to Oct 2019 https://www.similarweb.com/
[38] Number of monthly visits is based on report over the period from Aug 2019 to Oct 2019 https://www.similarweb.com/
[39] Ranking is accessed through https://www.alexa.com/siteinfo in Nov/15/2019.
[40] Ranking is based on daily trade volume, reported by CoinmarketCap https://coinmarketcap.com/ ,daily averaged during the sample period.

UT8	$\geqslant$ 5 year	24646	633	1.224	3636	14.5
UT9	$\geqslant$ 5 year	975	384	2.146	7682	95.6
UT10	$\geqslant$ 5 year	18452	517.5	1.449	5231	30.0
Panel U Unregulated Tier-2 exchanges						
U1	$<$ 2year	7805	17322	0.032	81142	29.9
U2	$<$ 2year	30997	N/A	0.260	3684	19.0
U3	2year $\leqslant$ A$<$5year	3464	4926.5	0.096	19860	16.1
U4	2year $\leqslant$ A$<$5year	50944	2594	0.234	30210	10.2
U5	2year $\leqslant$ A$<$5year	14534	5928.5	0.031	363745	46.6
U6	2year $\leqslant$ A$<$5year	52741	6735	0.092	6422	16.0
U7	$<$ 2year	34624	2770	0.265	6306	11.9
U8	$<$ 2year	21848	1818.5	0.092	100223	15.0
U9	2year $\leqslant$ A$<$5year	52	961.5	0.919	37634	90.0
U10	$<$ 2year	2756	11567	0.007	1684659	63.6
U11	2year $\leqslant$ A$<$5year	32305	3403.5	0.190	1714	16.8
U12	$<$ 2year	16035	3243	0.313	22780	30.8
U13	$<$ 2year	2612	2316.5	0.342	28739	30.4
U14	2year $\leqslant$ A$<$5year	16668	10350.5	0.032	53000	21.3
U15	$<$ 2year	23525	3061.5	0.188	1858	16.2
U16	$\geqslant$ 5 year	2013	1096.5	1.065	29808	73.7

Table 2 Chi-Squared test for conformity with Benford s law

Table 2 presents the Pearson's Chi-squared statistics and examine whether trade-size distributions of exchanges are consistent with distribution predicted by Benford's law. Distributions of four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Regulated exchanges are those which are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank. χ^2 statistics and *p*-value are reported in the table. ***, ** and * represents the statistical significant level at 1%, 5% and 1%, respectively.

Exchange Code	BTC/USD		ETH/USD		LTC/USD		XRP/USD	
	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value
Panel A Regulated exchanges								
R1	1.647	0.990	1.639	0.990	4.905	0.768	11.487	0.176
R2	2.736	0.950	2.767	0.948	3.218	0.920	2.189	0.975
R3	3.304	0.914	0.698	1.000	1.969	0.982		
Panel B Unregulated Tier-1 exchanges								
UT1	2.495	0.962	4.113	0.847	4.645	0.795	7.205	0.515
UT2	1.464	0.993	2.620	0.956	6.117	0.634	0.748	0.999
UT3	29.501***	0.000	5.349	0.720	7.157	0.520	47.121***	0.000
UT4	6.329	0.610	3.833	0.872	7.641	0.469	1.482	0.993
UT5	6.832	0.555	3.104	0.928	1.094	0.998	0.468	1.000
UT6	5.969	0.651	4.100	0.848	7.386	0.496	7.790	0.454
UT7	17.223**	0.028	4.823	0.776			3.644	0.888
UT8	2.601	0.957	1.956	0.982	3.724	0.881	4.230	0.836
UT9	3.228	0.919	7.886	0.445	2.454	0.964	14.219*	0.076
UT10	2.815	0.945	0.069	1.000	0.813	0.999	0.541	1.000
Panel C Unregulated Tier-2 exchanges								
U1	0.548	1.000	0.949	0.999				
U2	24.261***	0.002	16.677**	0.034	6.505	0.591	4.371	0.822
U3	4.660	0.793	19.569**	0.012	3.396	0.907	4.490	0.810
U4	1.360	0.995	2.468	0.963	0.673	1.000	0.723	0.999
U5	50.614***	0.000	8.254	0.409	124.881***	0.000	39.69***	0.000
U6	0.399	1.000	0.064	1.000				
U7	5.088	0.748	23.086***	0.003	60.516***	0.000	15.300*	0.054
U8	114.788***	0.000	141.768***	0.000	31.068***	0.000	57.021***	0.000
U9	63.022***	0.000	122.298***	0.000			71.949***	0.000
U10	10.771	0.215	4.662	0.793	12.325	0.137	26.135***	0.001
U11	2.430	0.965	7.140	0.522	4.115	0.847	7.602	0.473
U12	0.544	1.000	0.122	1.000	1.042	0.998	14.676*	0.066
U13	1.157	0.997	2.583	0.958	11.614	0.169	4.815	0.777
U14	0.678	1.000	23.351***	0.003	109.944***	0.000	26.835***	0.001
U15	2.240	0.973	0.536	1.000	0.703	1.000	2.249	0.972
U16	1.695	0.989	0.924	0.999	1.317	0.995	0.577	1.000

Table 3 Students' *t*-test analysis for trade-size clustering

Table 3 reports t-test analysis for trade size-clustering effect and examine whether trade frequencies at rounded size is larger than frequencies in adjacent trade size. Four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. For each cryptocurrency, observation ranges center on rounded sizes: multiples of 100 units with a radius 50 units (100X-50, 100X+50), and multiples of 500 units with a radius 100 units (500Y-100, 500Y+100). A positive difference indicates that frequency at rounded size is higher than frequency at adjacent unrounded size, therefore leading to trade clustering at rounded trade size. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank. Differences and t-statistics are reported in the table. ***, ** and * represents the positive and statistical significant level at 1%, 5% and 1%, respectively.

Observation range: Multiples of 100 units (100X-50, 100+50)

Exchange Code	BTC/USD		ETH/USD		LTC/USD		XRP/USD	
	Difference	*t* statistics	Difference	*t* statistics	Difference	*t* statistics	Difference	*t* statistics
Panel A Regulated exchanges								
R1	0.0909***	14.4897	0.1121***	12.2797	0.1598***	10.7672	0.0628***	6.7260
R2	0.0890***	14.8747	0.1352***	15.6466	0.1085***	8.9447	0.0316***	2.9547
R3	0.1253***	13.6550	0.1186	9.7127	0.2028***	8.2836		
Panel B Unregulated Tier-1 exchanges								
UT1	0.1884***	16.9926	0.2259***	20.7396	0.1787***	9.3103	0.0048	0.5403
UT2	0.0255*	1.9259	0.0387**	2.3273	0.0645***	2.9432	0.0757***	3.9519
UT3	0.1003***	12.6538	0.0782***	8.6550	0.1095***	6.6961	0.0761***	5.6808
UT4	0.0052	1.0732	-0.0024	-0.5683	0.0044	0.6442	-0.0050	-0.5559
UT5	0.1280***	16.8946	0.0829***	14.4424	0.1040***	8.0033	0.0104	1.1159
UT6	-0.0151	-2.6682	-0.0011	-0.0807	-0.0028	-0.0889	-0.0136	-1.3791
UT7	0.0884***	6.8536	0.0565***	3.6850			0.1319***	6.4976
UT8	0.0815***	12.6202	0.0665***	10.6136	0.0470***	5.2893	0.0086	0.9026
UT9	0.0839***	10.1919	0.0604***	5.7821	0.1006***	4.0182	0.0543**	2.5696
UT10	-0.0127	-4.1187	-0.0161	-18.6353	-0.0299	-9.1732	-0.0197	-16.2060
Panel C Unregulated Tier-2 exchanges								
U1	-0.0159	-86.2084	-0.0220	-7.3739				
U2	-0.0149	-24.7328	-0.0136	-12.2973	-0.0168	-27.7010	-0.0167	-34.6752
U3	0.0299***	7.1098	0.0288***	3.6873	-0.0021	-0.1310	-0.0829	-2.2636
U4	-0.0079	-5.6289	-0.0153	-5.4152	-0.0116	-2.6007	-0.0077	-1.0193
U5	0.0729***	6.5729	-0.0268	-7.2791	-0.0146	-13.8437	-0.0140	-11.1992
U6	-0.0203	-33.1736	-0.0223	-52.8755				
U7	0.0194*	1.9516	0.0959***	9.0186	0.0581***	9.9824	-0.0172	-15.2209
U8	-0.0009	-0.3406	0.0354***	6.5518	-0.0054	-0.8043	-0.0078	-1.2069
U9	0.1055**	2.3127	0.0323	1.0382			-0.0224	-0.4503
U10	-0.0035	-5.6219	-0.0151	-11.5489	-0.0164	-12.7298	-0.0147	-22.7750
U11	0.2588***	20.2791	0.1225***	31.4661	0.1107***	15.2575	-0.0167	-16.1563
U12	-0.0149	-13.1637	-0.0139	-15.8460	-0.0211	-15.3038	-0.0349	-3.1582
U13	0.0339***	3.4112	0.0608***	8.3164	0.0944***	5.6619	0.0831***	6.5028
U14	-0.0317	-22.4364	-0.0212	-33.1229	-0.0365	-16.1747	-0.0329	-2.1494
U15	-0.0149	-8.2660	-0.0150	-8.7648	-0.0176	-35.6842	-0.0167	-30.5822
U16	0.2430***	20.5746	0.0194**	2.3537	0.0184*	1.7527	0.0044	0.3330

Observation range: Multiples of 500 units (500X-100, 500X +100)

Exchange Code	BTC/USD		ETH/USD		LTC/USD		XRP/USD	
	Difference	*t* statistics	Difference	*t* statistics	Difference	*t* statistics	Difference	*t* statistics
Panel A Regulated exchanges								
R1	0.2028***	15.1934	0.2706***	15.5333	0.2481***	7.9035	0.1655***	7.8488
R2	0.1946***	16.7576	0.2897***	18.5028	0.2064***	9.9654	0.1366***	5.8932
R3	0.2662***	13.1447	0.3096***	13.3761	0.3313***	7.7499		
Panel B Unregulated Tier-1 exchanges								
UT1	0.3542***	25.2234	0.3911***	35.1598	0.3929***	16.1712	0.0826***	3.5290
UT2	0.0964***	2.9995	0.1018***	2.8984	0.1142	1.6911	0.1372***	3.5442
UT3	0.2207***	13.6262	0.1931***	12.2024	0.2363***	7.8381	0.1974***	6.0042
UT4	0.0386***	2.9776	0.0329***	3.5719	0.0393**	2.0862	0.0347	1.6019
UT5	0.2574***	24.0095	0.1465***	19.7693	0.1975***	10.8497	0.0590***	3.0175
UT6	-0.0184	-2.3417	0.0244	0.8890	0.0690	0.9597	-0.0304	-1.4275
UT7	0.1847***	5.6031	0.1709***	4.9384			0.2469***	5.7462
UT8	0.1385***	16.4183	0.1051***	13.0114	0.0773***	5.6468	0.0346**	2.0117
UT9	0.1634***	6.3122	0.1587***	7.0994	0.2388***	4.5180	0.0958***	2.7676
UT10	-0.0096	-2.0254	-0.0088	-6.0407	-0.0287	-3.6788	-0.0126	-7.4568
Panel C Unregulated Tier-2 exchanges								
U1	-0.0078	-45.0622	-0.0144	-2.5713				
U2	-0.0072	-18.6149	-0.0022	-0.5960	-0.0093	-10.8383	-0.0087	-12.0355
U3	0.0074	1.1223	0.0409**	2.3658	-0.0554	-1.1327	-0.0695	-0.8426
U4	-0.0049	-3.5093	-0.0010	-0.1424	0.0062	0.4515	-0.0014	-0.0960
U5	-0.0089	-3.2615	-0.0139	-4.0280	-0.0064	-3.8899	-0.0065	-8.5313
U6	-0.0140	-11.8150	-0.0123	-17.5248				
U7	0.0787**	2.0781	0.2463***	15.4851	0.0181*	2.0080	-0.0086	-7.7077
U8	0.0057	1.3326	0.0297***	3.4984	-0.0002	-0.0224	0.0034	0.4150
U9	0.1821**	2.8803	0.0701	1.1542			0.0591	0.6019
U10	-0.0016	-6.4914	-0.0069	-16.3422	-0.0127			
U11	0.3694***	11.1562	0.0610***	9.8832	0.0619***	5.5217	-0.0079	-13.6864
U12	-0.0012	-0.7434	-0.0079	-12.1335	-0.0120	-8.1838		
U13	0.1498***	5.9352	0.0980***	6.7196	0.0535***	2.8451	0.1548***	6.9233
U14	-0.0203	-11.9796	-0.0125	-13.5750	-0.0220	-9.6109	0.0009	0.1201
U15	-0.0043	-0.6221	-0.0014	-0.1846	-0.0092	-10.5385	-0.0080	-15.6314
U16	0.2192***	8.5892	0.0797***	4.4893	0.0514**	2.4985	0.0361	1.4418

Table 4 Power law exponents

Table 4 presents the scaling exponents for each crypto exchange in power law models. Two scaling parameters $\hat{\alpha}_{OLS}$ and $\hat{\alpha}_{Hill}$ are included and estimated by Ordinary Least Square (OLS) and Maximum Likelihood Estimation (MLE), respectively.[41] Histogram of probability density on log-log scale is adopted in the estimation. Four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. We also check whether the estimated parameters are in the Pareto–Lévy range (1<α<2) and mark Y if both exponents lie in the Pareto–Lévy range. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

Exchange Code	BTC/USD			ETH/USD			LTC/USD			XRP/USD		
	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{Hill}$	Pareto–Lévy (1<α<2)	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{Hill}$	Pareto–Lévy (1<α<2)	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{Hill}$	Pareto–Lévy (1<α<2)	$\hat{\alpha}_{OLS}$	$\hat{\alpha}_{Hill}$	Pareto–Lévy (1<α<2)
Panel A Regulated exchanges												
R1	1.8063	1.2793	Y	1.6962	1.3741	Y	1.5097	1.8490	Y	1.7479	1.3383	Y
R2	1.7632	1.1914	Y	1.7447	1.3081	Y	1.8573	1.3085	Y	1.8094	1.2570	Y
R3	1.6680	1.2972	Y	1.7615	1.4252	Y	1.6731	1.8348	Y			
Panel B Unregulated Tier-1 exchanges												
UT1	1.6693	1.2091	Y	1.7947	1.4357	Y	1.8357	1.4108	Y	1.9600	1.4304	Y
UT2	1.9105	1.6710	Y	1.5816	1.8802	Y	1.8065	1.4972	Y	1.7984	1.7223	Y
UT3	1.6804	1.2766	Y	1.7187	1.4246	Y	1.8154	1.3974	Y	1.9475	1.4303	Y
UT4	0.6196	0.6628	N	0.7851	0.7902	N	0.6921	0.8791	N	0.5522	0.8034	N
UT5	1.7495	1.0885	Y	1.8415	1.5052	Y	1.8712	1.4472	Y	1.9658	1.6508	Y
UT6	3.3248	1.6564	N	3.0142	1.6087	N	4.5630	5.8653	N	5.9756	5.5786	N
UT7	1.4063	0.9051	N	1.4944	1.3583	Y				1.2818	1.2314	Y
UT8	1.6796	0.9487	N	1.6745	1.0196	Y	1.8628	1.3203	Y	1.8123	1.2121	Y
UT9	1.6291	1.0077	Y	1.6146	1.8164	Y	1.6616	1.4281	Y	1.8044	1.4699	Y
UT10	1.4785	1.0953	Y	1.8410	1.4173	Y	1.5462	0.9319	N	1.6339	1.1936	Y
Panel C Unregulated Tier-2 exchanges												
U1	1.3329	2.7600	N	3.3447	3.9405	N						
U2	5.1967	7.1552	N	10.4280	7.0760	N	1.7393	2.0456	N	2.1936	1.4686	N
U3	2.3735	2.7024	N	2.0350	1.5458	N	2.0136	4.0051	N	2.2020	4.4523	N
U4	4.5461	2.7244	N	4.7157	3.5726	N	7.1649	4.1373	N	6.3558	4.1571	N

[41] We apply the probability density function to estimate the scaling exponents 1+ α.

U5	2.2687	1.7014	N	4.3669	1.7726	N	0.6414	1.2986	N	8.6893	4.8634	N
U6	1.7603	1.6380	Y	1.9977	1.6221	Y						
U7	7.6601	7.0626	N	3.5977	11.4442	N	14.8153	11.7058	N	12.4393	6.8617	N
U8	1.0199	0.9517	N	1.1573	0.8744	N	1.2411	0.7654	N	0.6564	0.6497	N
U9	1.3699	3.7699	N	1.5201	3.0867	N				1.4860	6.3731	N
U10	4.2924	7.5782	N	7.3844	7.9664	N	5.0487	8.8024	N	10.6973	13.8628	N
U11	5.8287	6.3843	N	3.6391	5.9611	N	3.6763	4.8774	N	7.1159	5.0273	N
U12	2.8543	1.7280	N	1.9257	1.8797	Y	1.5719	1.2255	Y	1.8310	2.6912	N
U13	1.5094	1.0223	Y	1.6690	1.1908	Y	1.4789	1.1925	Y	1.4344	1.1797	Y
U14	0.7176	1.2605	N	2.0313	1.2371	N	1.0767	1.0560	Y	6.5510	10.5244	N
U15	1.5367	1.0376	Y	1.6176	1.1166	Y	1.6789	1.1291	Y	1.5482	1.0010	Y
U16	2.0480	1.6306	N	1.9248	1.9540	Y	2.1730	2.4296	N	2.1748	2.0735	N

Table 5 Chi-Squared test for level of roundness

Table 5 presents the Pearson's Chi-squared statistics and compares the frequency distribution in different level of roundness between regulated and unregulated exchanges. Distributions of four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. The level of roundness is negatively associated with the number of effective decimal figures. We next examine the frequency distribution over level of roundness, measured by place value of last non-zero digit. The average distribution of regulated exchanges is used as a benchmark and compared with unregulated exchanges. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank. χ^2 statistics and p-value are reported in the Table. ***, ** and * represents the statistical significant level at 1%, 5% and 10%, respectively.

Exchange Code	BTC/USD		ETH/USD		LTC/USD		XRP/USD	
	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value
Panel A Regulated exchanges								
R1	1.698	0.945	7.755	0.257	1.531	0.909	4.283	0.233
R2	2.608	0.856	5.140	0.526	4.724	0.451	4.283	0.233
R3	1.815	0.936	2.903	0.821	3.292	0.655		
Panel B Unregulated Tier-1 exchanges								
UT1	9.545	0.145	15.013**	0.020	12.18**	0.032	11.993***	0.007
UT2	3.100	0.796	11.455*	0.075	9.222	0.101	13.387***	0.004
UT3	92.104***	0.000	8.086	0.232	5.616	0.345	51.094***	0.000
UT4	17.224***	0.008	13.387**	0.037	7.547	0.183	11.393***	0.010
UT5	115.48***	0.000	11.01*	0.088	14.311**	0.014	9.5**	0.023
UT6	7.909	0.245	17.469***	0.008	24.886***	0.000	16.603***	0.001
UT7	182.435***	0.000	16.518**	0.011			49.766***	0.000
UT8	4.384	0.625	15.649**	0.016	19.46***	0.002	12.18***	0.007
UT9	3.247	0.777	5.427	0.490	11.906**	0.036	14.268***	0.003
UT10	1461.8***	0.000	692.292***	0.000	21.797***	0.001	18.032***	0.000
Panel C Unregulated Tier-2 exchanges								
U1	18.774***	0.005	32.402***	0.000				
U2	60.923***	0.000	62.726***	0.000	28.101***	0.000	19.651***	0.000
U3	828.828***	0.000	85.86***	0.000	22.242***	0.000	19.593***	0.000
U4	1670.819***	0.000	31.158***	0.000	32.097***	0.000	19.747***	0.000
U5	1668.236***	0.000	20.761***	0.002	27.753***	0.000	19.109***	0.000
U6	1639.493***	0.000	24.944***	0.000				
U7	9.569	0.144	15.481**	0.017	18.705***	0.002	19.688***	0.000
U8	740.835***	0.000	157.443***	0.000	86.741***	0.000	18.59***	0.000
U9	15.455**	0.017	26.838***	0.000			19.182***	0.000
U10	1719.65***	0.000	23.694***	0.001	32.242***	0.000	19.796***	0.000
U11	439.322***	0.000	101.26***	0.000	14.106**	0.015	19.458***	0.000
U12	18.605***	0.005	28.754***	0.000	22.785***	0.000	19.768***	0.000
U13	26.08***	0.000	130.687***	0.000	41.623***	0.000	34.596***	0.000
U14	1310.242***	0.000	34.176***	0.000	30.144***	0.000	19.728***	0.000
U15	1546.727***	0.000	23.247***	0.001	29.609***	0.000	19.592***	0.000
U16	535.379***	0.000	55.367***	0.000	13.247**	0.021	15.288***	0.002

Table 6 Estimation of Wash Trading Sizes

Table 6 reports the pooled regression results of the log of non-round volumes on the log of round volumes for all regulated exchanges with a weekly frequency.

$$\ln(V_{Unround_{it}}) = \alpha + \beta * \ln(V_{Round_{it}}) + \mu_{it}$$

where $\ln(V_{Round_{it}})$ and $\ln(V_{Unround_{it}})$ are the logarithms of rounded trade volume and unrounded trade volumes, respectively, of exchange i at week t. We categorize trading volumes into rounded and unrounded ones by checking if the mantissa of a certain transaction volume is less than 100 base units or not. Wash trades in unregulated exchanges are the quantity of unrounded volume deducted by the non-wash volume estimated using rounded trading volumes.

	Wash Volume Percentage	
	Average	Standard Deviation
Unregulated exchanges	69.72%	29.71%
Unregulated Tier-1 exchanges	52.52%	29.41%
Unregulated Tier-2 exchanges	80.48%	25.13%

Exchange Code	Wash Volume Percentage
Panel A Unregulated Tier-1 Exchanges	
UT1	51.76%
UT2	51.73%
UT3	1.12%
UT4	92.60%
UT5	44.87%
UT6	66.3%
UT7	18.95%
UT8	66.12%
UT9	37.49%
UT10	94.31%
Panel B Unregulated Tier-2 Exchanges	
U1	99.99%
U2	98.30%
U3	72.72%
U4	95.50%
U5	89.71%
U6	98.13%
U7	77.20%
U8	77.09%
U9	81.12%
U10	98.45%
U11	21.48%
U12	98.08%
U13	65.42%
U14	96.78%
U15	94.36%
U16	23.27%

Table 7 Relationship between Failure Rates and Fraction of Wash Trades

Relation between failure rates and fraction of wash trade is examined with regression analysis. *t*-statistics are reported in the brackets. ***, ** and * represents the statistical significant level at 10%, 5% and 1%, respectively.

Fraction of wash trade	Unregulated exchange
Percentage of Failed Tests	0.597***
	(4.99)
Constant	0.412***
	(4.54)
Observations	26
Adjusted R2	35.2%

Figure 1First digit distribution with Benford s law

Figure 1 displays the first digit distribution and compared with distribution predicted by Benford s law. Distributions of four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Y axis represents the frequency of trades with a certain digit in the initial position. Panel R, Panel UT, and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

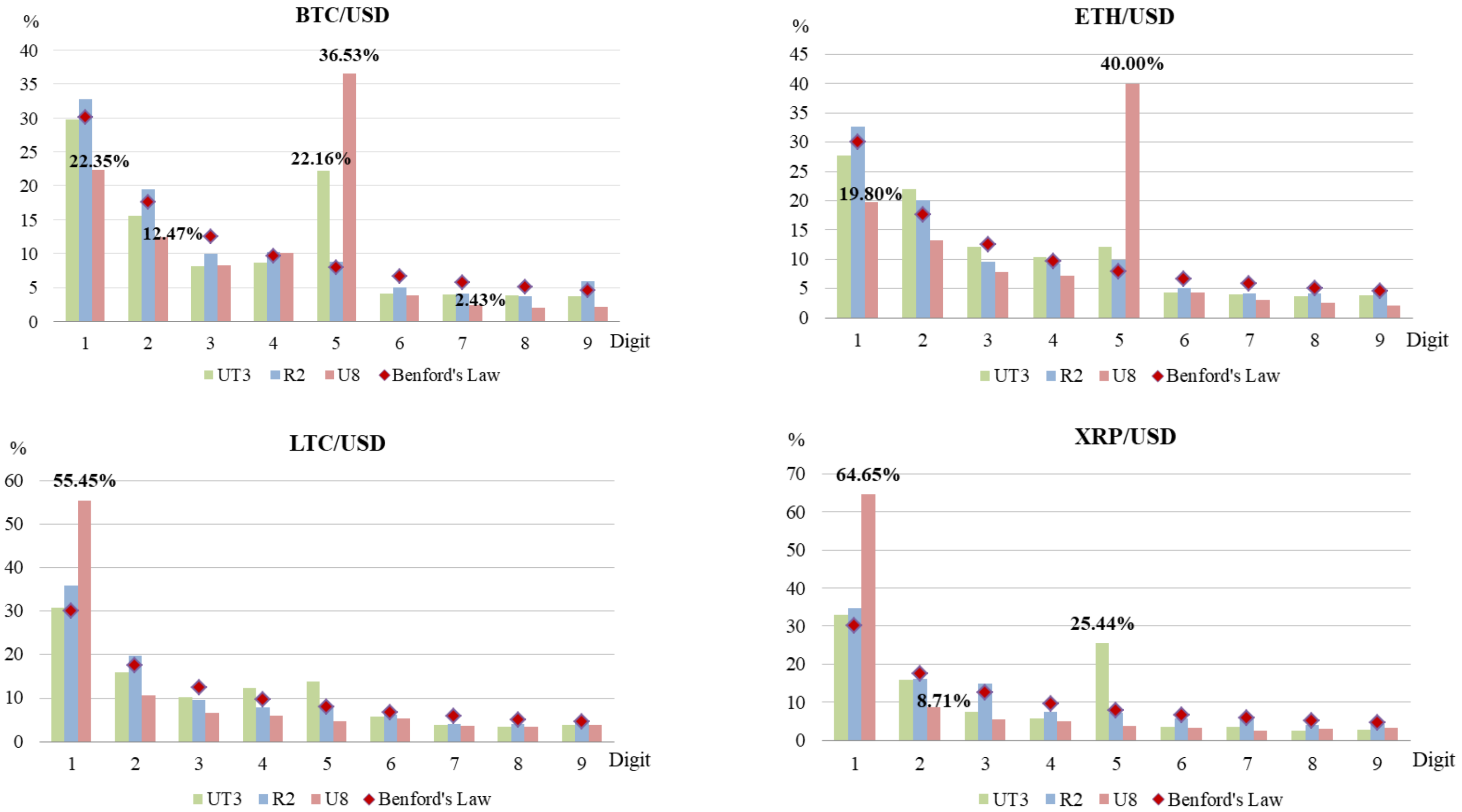
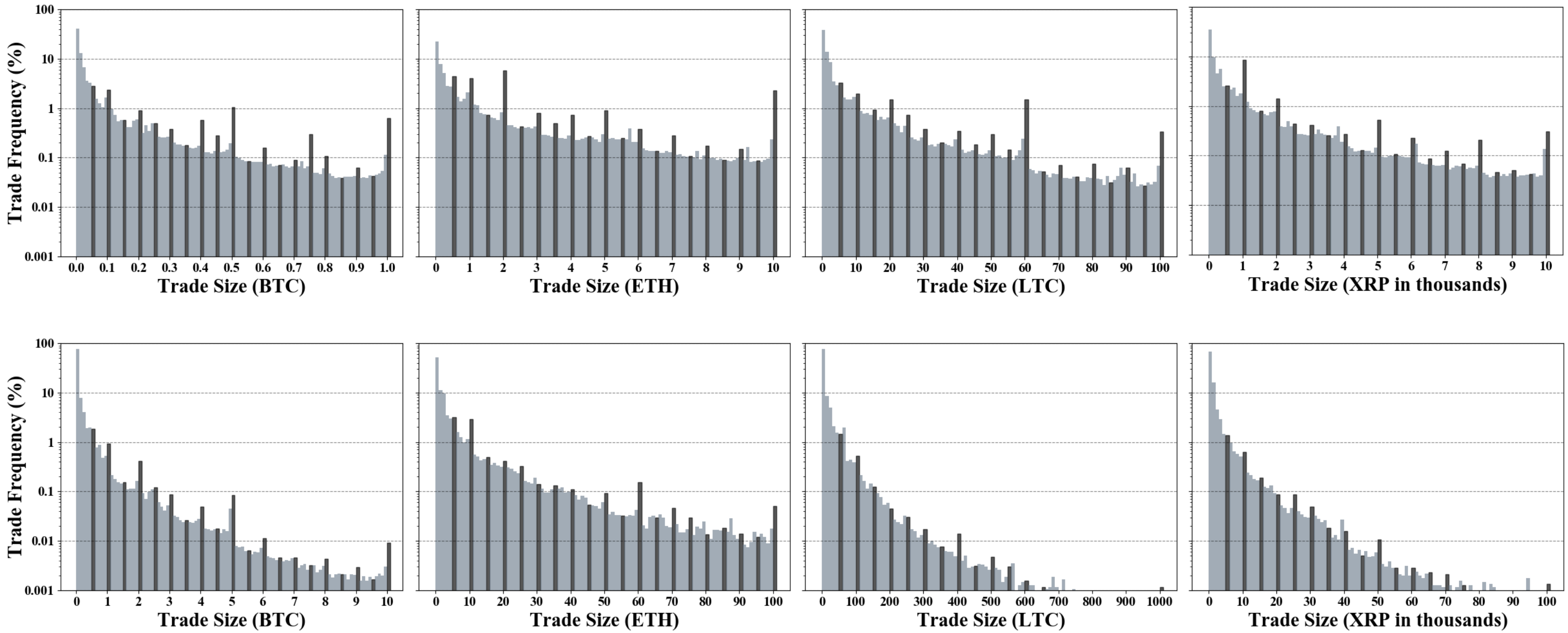


Figure 2 Trade lot size clustering

Figure 2 depicts the size clustering effect in trade size distributions for four trading pairs, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Histograms for four trading pairs are depicted over ranges of trade size, 0-1BTC, 0-10BTC, 0-10 ETH, 0-100ETH, 0-100LTC, 0-1000LTC, 0-10,000XRP and 0-100,000XPR. In each observation range, we highlight every 5th and 10th bins to illustrate the clustering effect around rounded trade sizes Y-axis represents the frequency of trades on a log scale. Panel R, Panel UT, and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

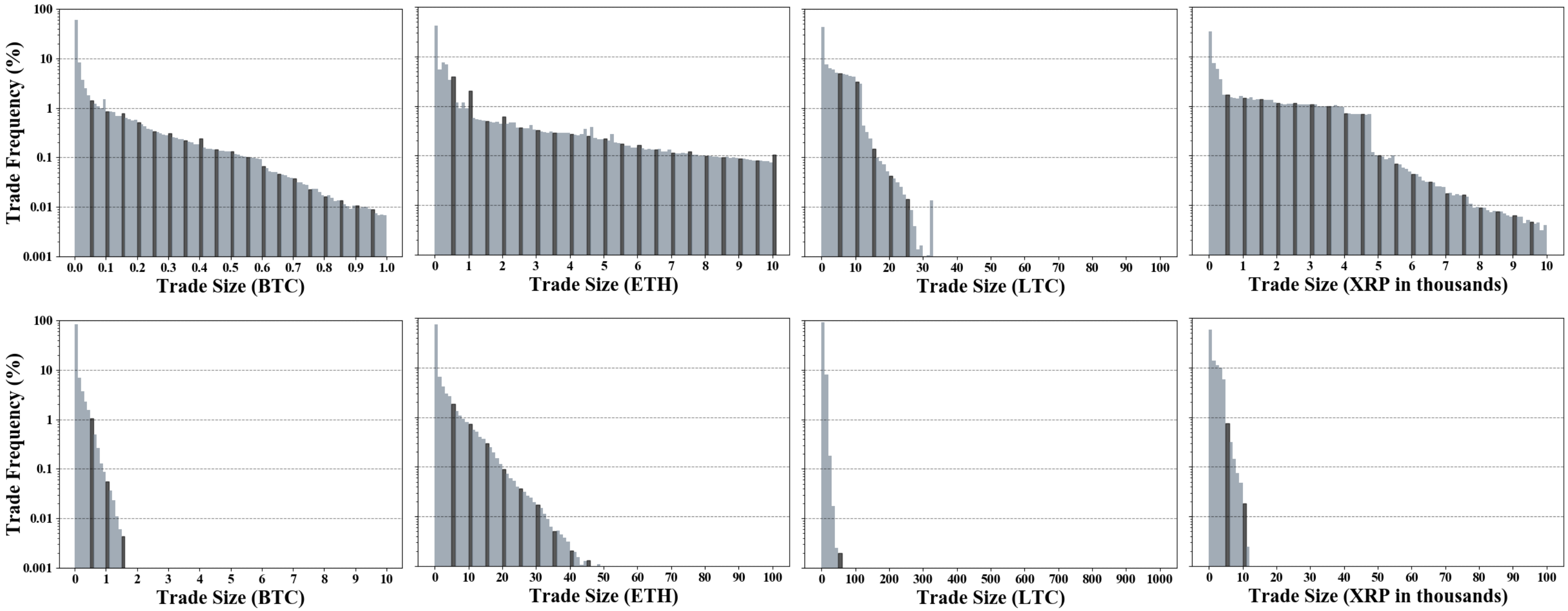
Panel R: Regulated exchanges

R2



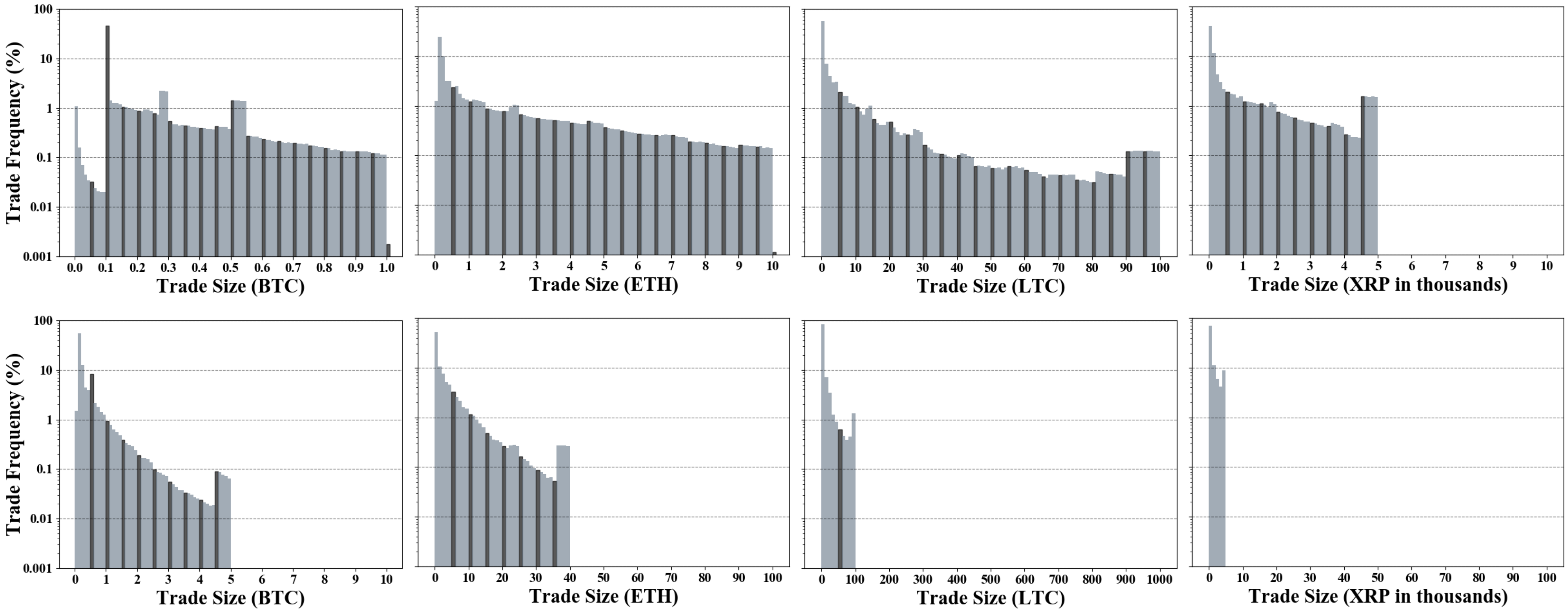
Panel UT: Unregulated Tier-1 exchanges

UT6

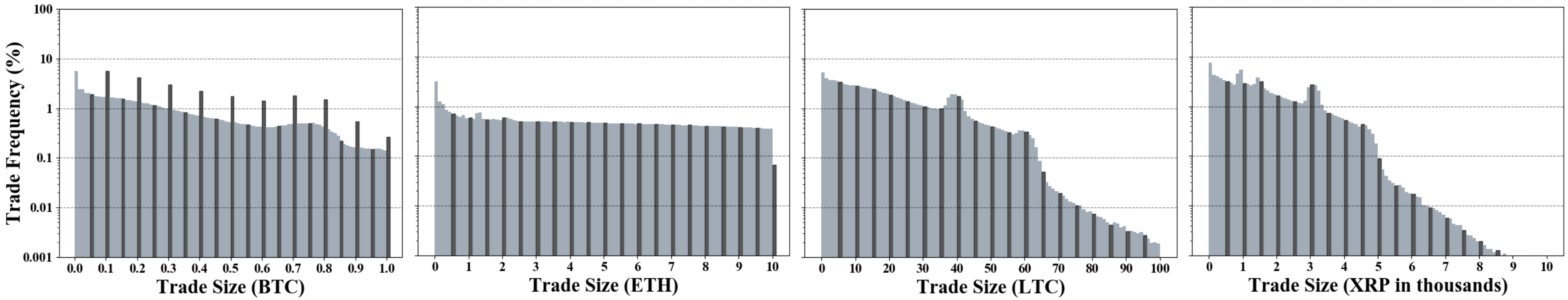


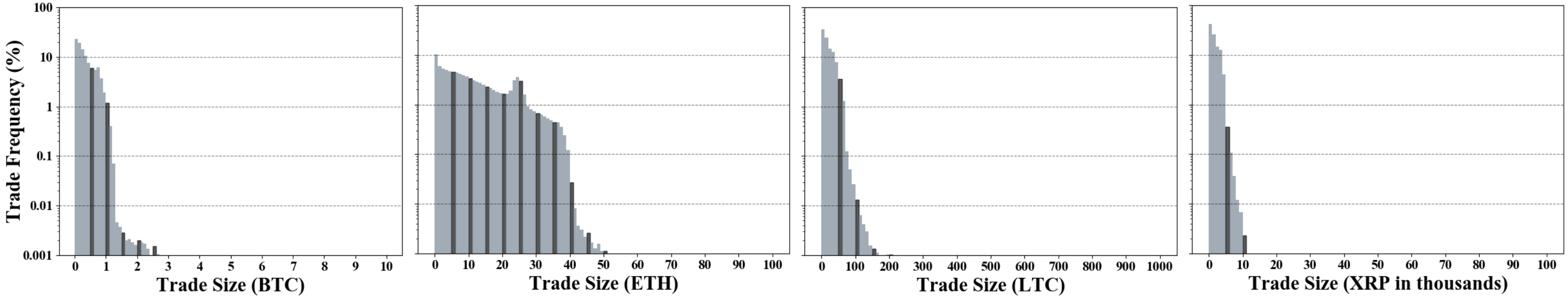
Panel U: Unregulated Tier-2 exchanges

U5



U11




Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U14

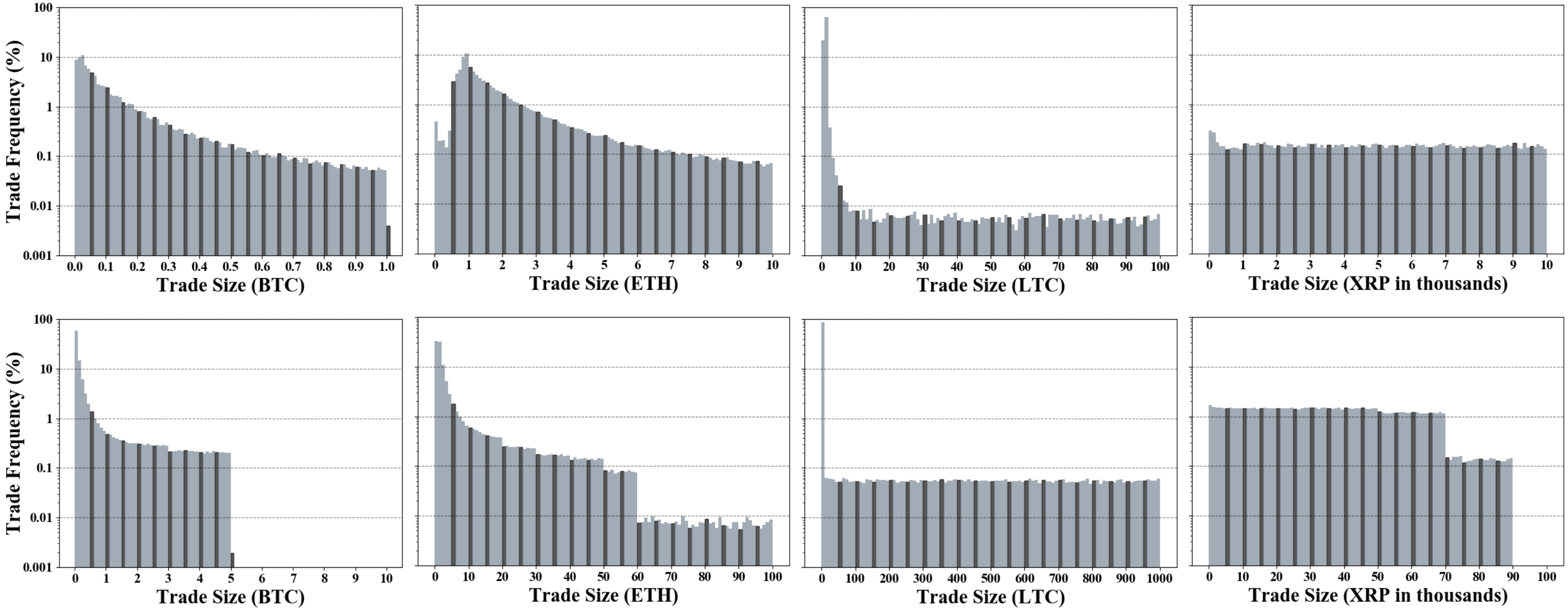

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

Figure 3 Calculation of Clustered Trade Frequency within an Observation Window

Trade frequency is calculated as the number of trades with size i over total trade numbers in the observation window. There is a much higher frequency at 200 base units than other places in the window.

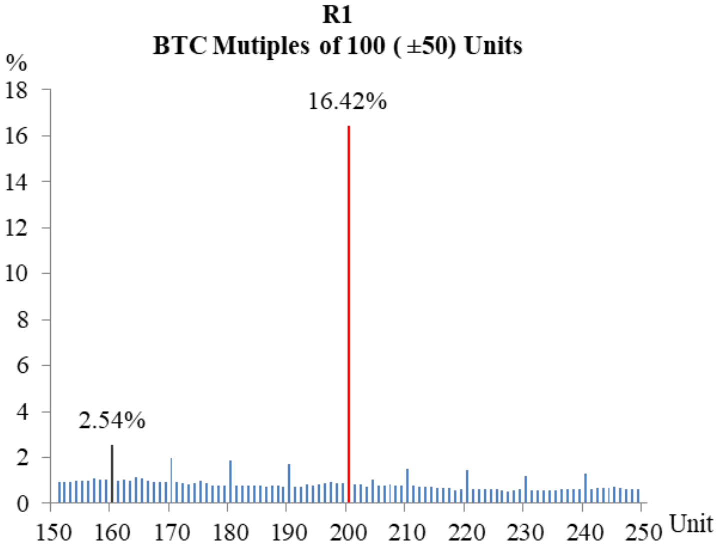
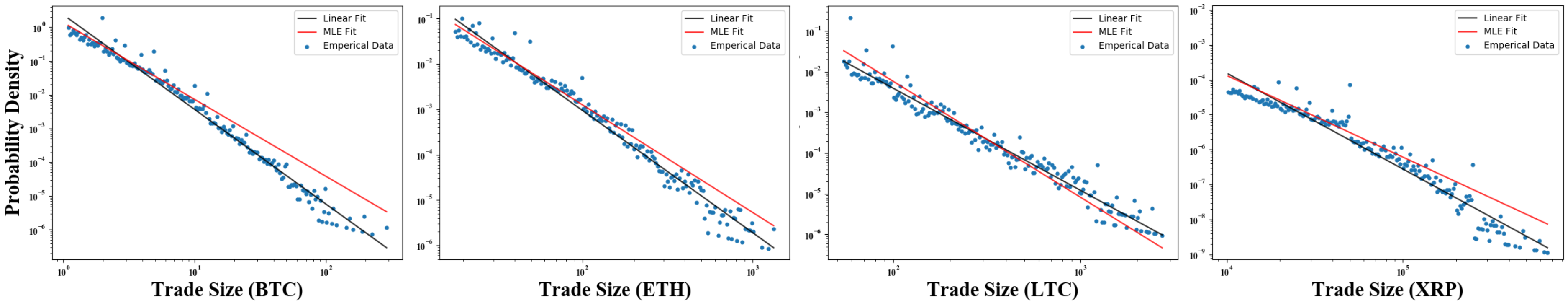


Figure 4 Tail Distribution and Fitted Power Law Line

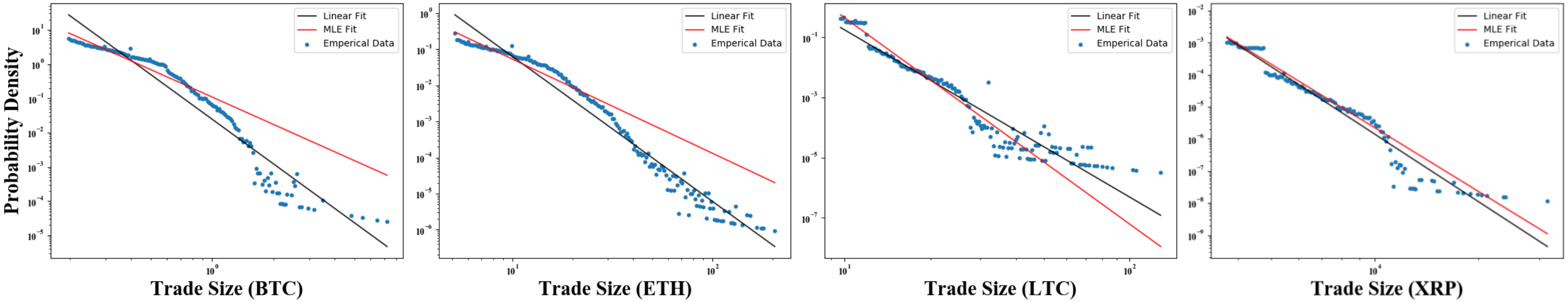
Figure 4 display tail distribution of trade size and fitted power-law line on log-log scale. For each crypto exchange, four trading pairs are presented, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Y-axis represents the frequency of trades on a log scale while X-axis represents the logarithm of trade size. Fitted power-law lines are estimated by Ordinary Least Square (OLS) and Maximum Likelihood Estimation (MLE), plotted as black solid lines and red lines, respectively. Blue dots represent empirical data points for trade size. Panel R, Panel UT, and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

Panel R: Regulated exchange

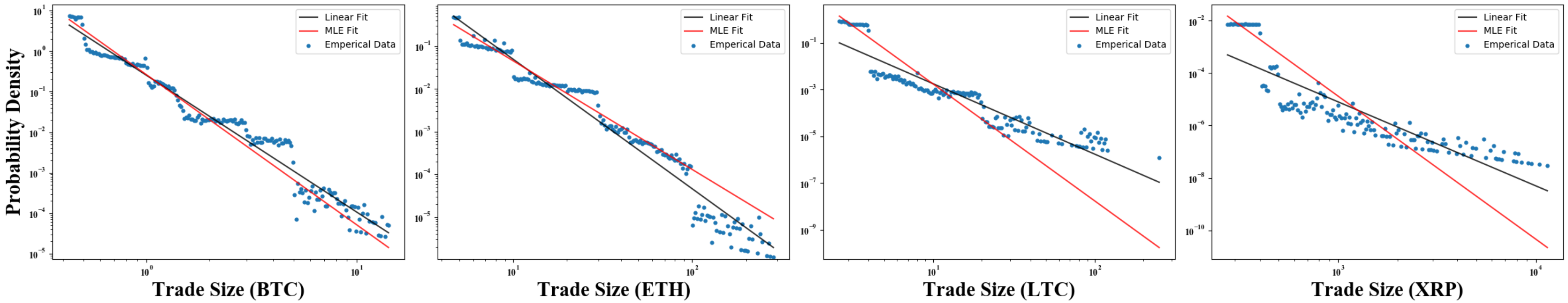
R1



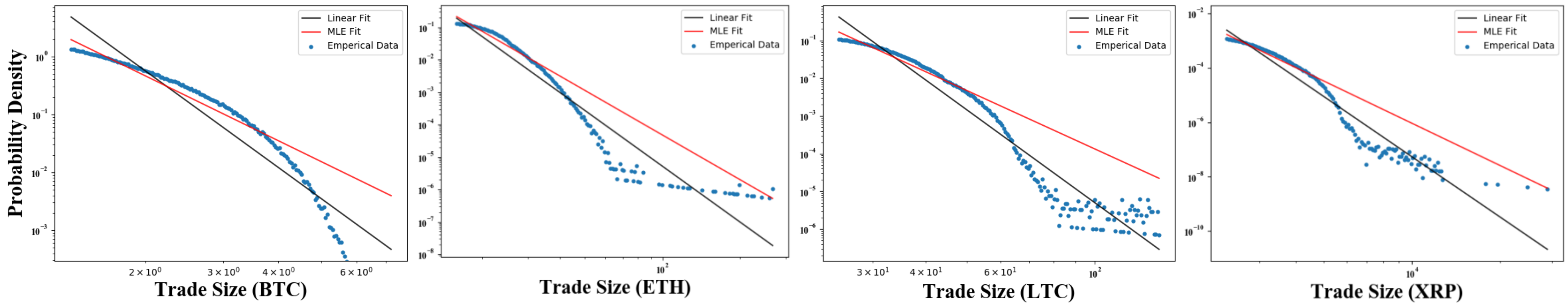
Panel UT: Unregulated Tier-1 exchanges
UT6



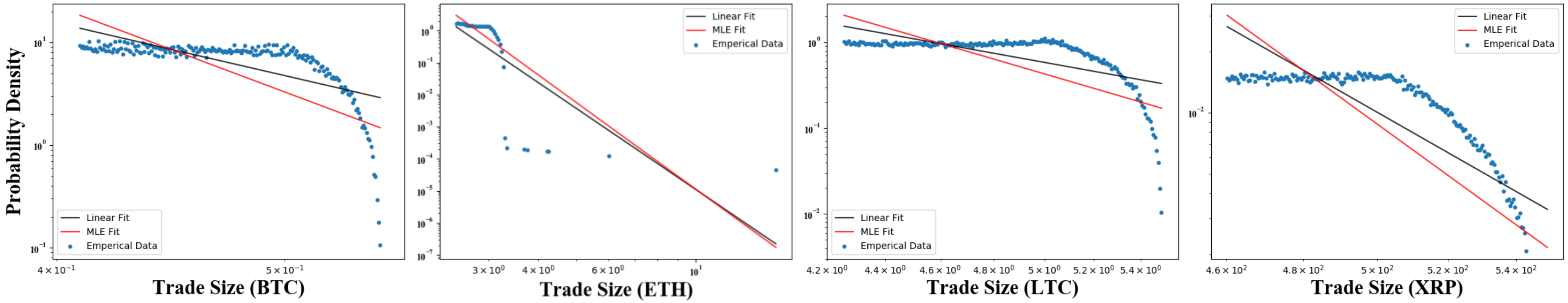
Panel U: Unregulated Tier-2 exchanges
U3



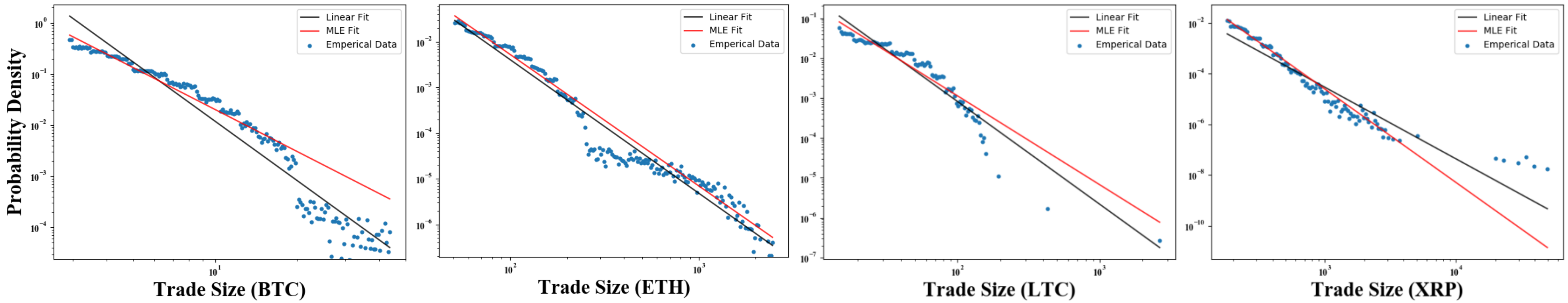
U4



U10


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U12


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U14

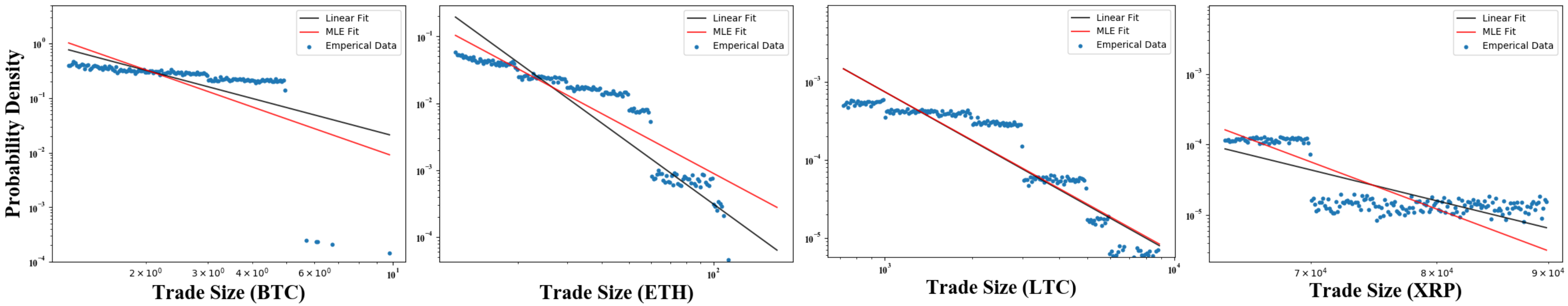
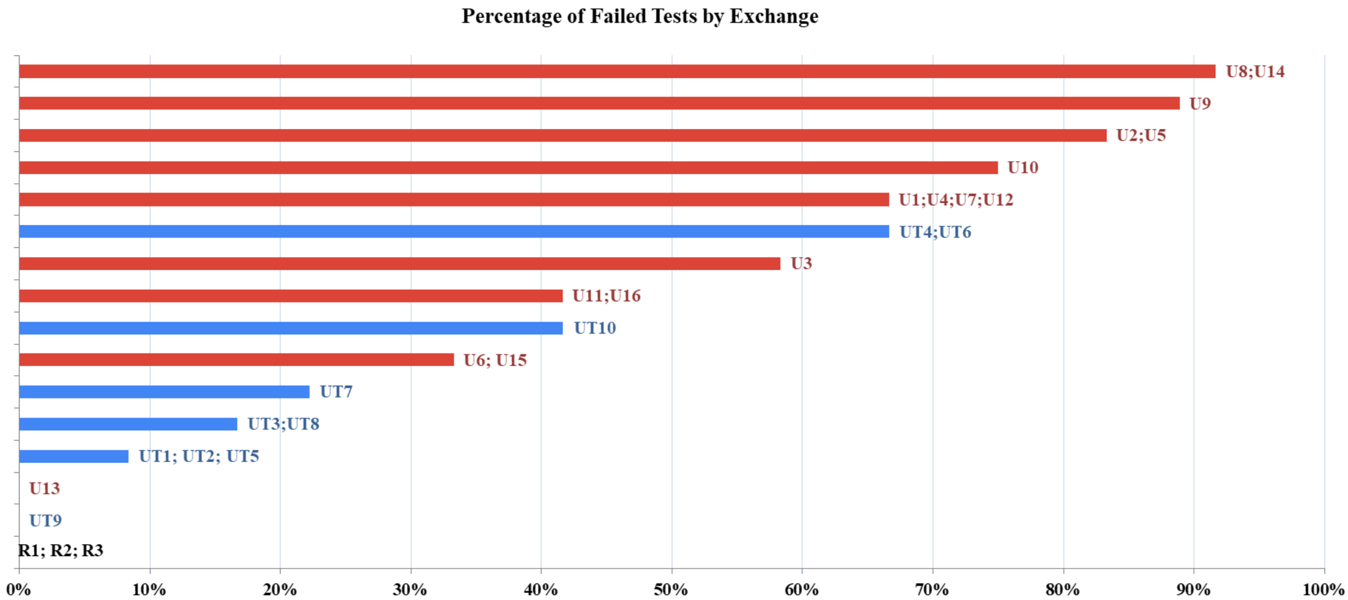

Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

Figure 5 Percentage of failed tests

Figure 5 presents percentage of failed tests for each crypto exchange. We summarize all statistical tests in Section 4 for each crypto exchange, including Chi-squared tests for Benford s Law, t-test for trade size clustering, and scaling exponents for power law. For each test, we report four cryptocurrency pairs, BTC, ETH, LTC and XRP. For each exchange, percentage of failed tests is calculated as the number of failed tests over the total number of tests across cryptocurrency (See details in Appendix F). Similarly, percentage of failed tests for each type of cryptocurrency is measured as the number of failed test at 5% significance level over the total number of tests in one type of cryptocurrency. In Chi-squared tests of first significant digits, "failure of test" is defined that exchange failed to conform to Benford s Law or trade pattern predicted by regulated exchanges, statistically significant at 5% or 1% level; "pass" otherwise. In *t*-test of size clustering, "failure of test" refers to that exchange do not show apparent size clustering at multiple of 100 units while "pass" represents noticeable clustering effect at 5% or 1% significance level. In estimation of power law exponents, "failure of test" refers to that scaling exponent either $\hat{\alpha}_{OLS}$or $\hat{\alpha}_{Hill}$ lie outside the Pareto–Lévy range (1, 2) and tail distribution does not show linear trend on log-log plot; "pass" otherwise.



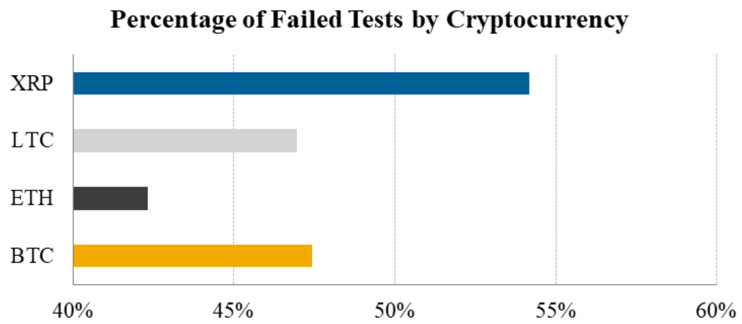
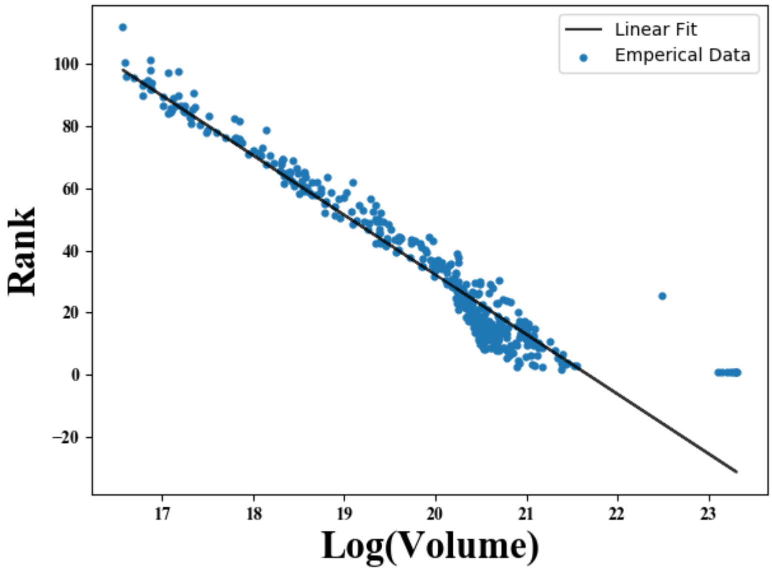


Figure 6

Figure 6 plots the quantitative relationship between trade volume (logarithm) and exchange rank. Data fitting is carried out with Ordinary Least Square (OLS) regression. The fitting line is shown in the figure, and estimated coefficients are reported below(t-statistics in brackets) with an adjusted R^2 of 93%.

$$\text{Exchange rank}_i = \underset{(78.01)}{416.269} - \underset{(-71.53)}{19.202} * \log(\text{Volume}_i) + \varepsilon_i$$



Appendix

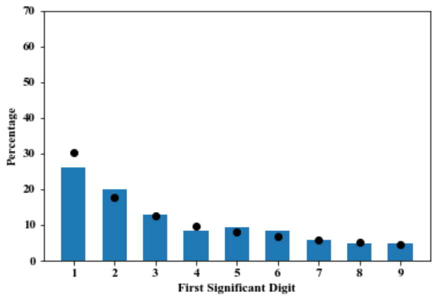
Appendix A First digit distribution of trade size

Appendix A displays the first digit distribution of trade size for each crypto exchange. Herein, we report the distribution of four trading pairs, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Distribution predicted by Benford s law is also plotted in each Figure. Y-axis represents frequency of trades with certain digit in initial position. Panel R, Panel UT and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

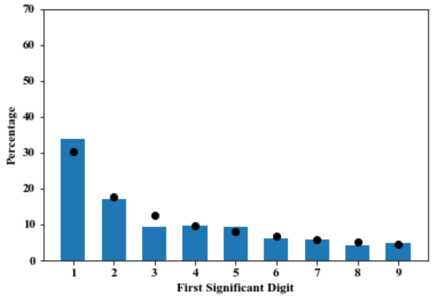
Panel R: Regulated exchange

R1

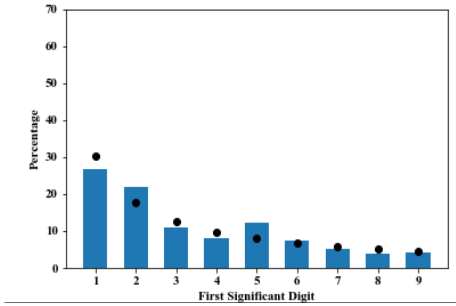
BTC/USD



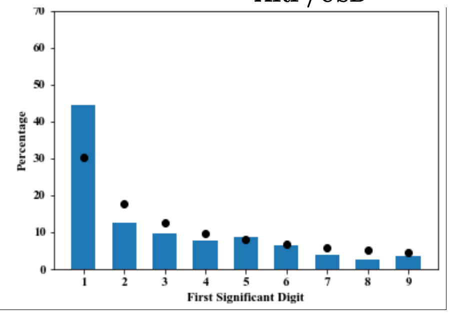
ETH/USD



LTC/USD

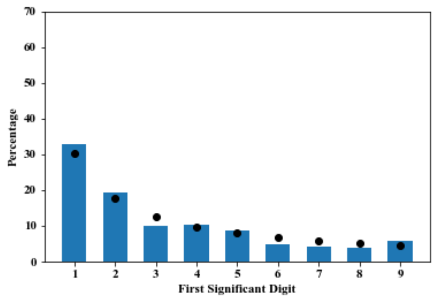


XRP/USD

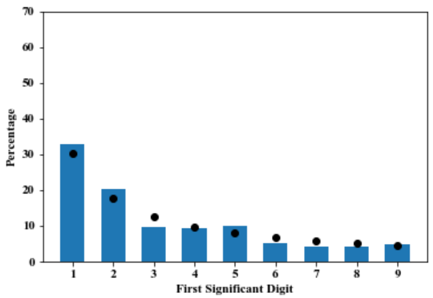


R2

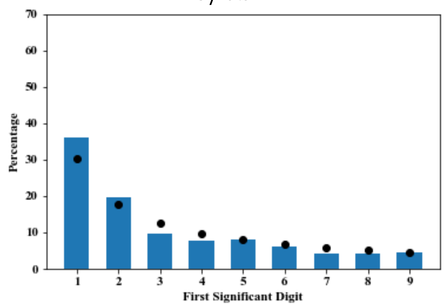
BTC/USD



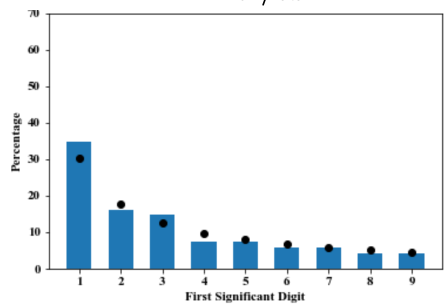
ETH/USD



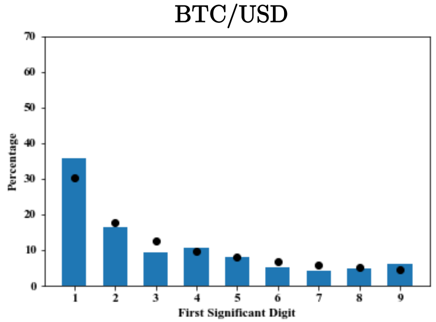
LTC/USD

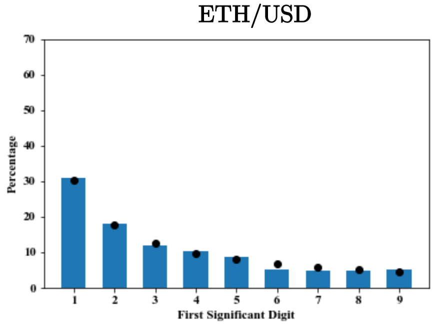


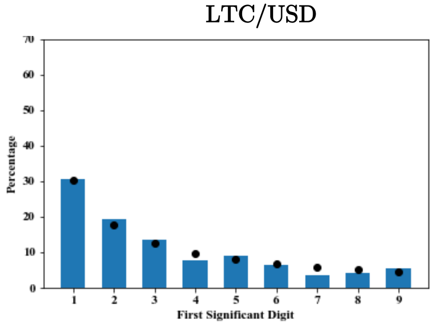
XRP/USD



R3

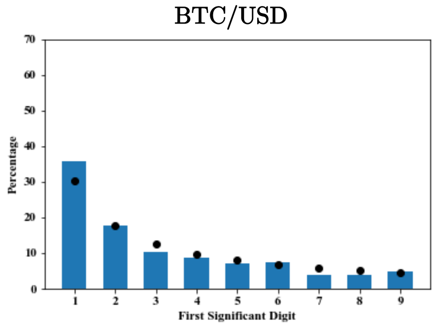


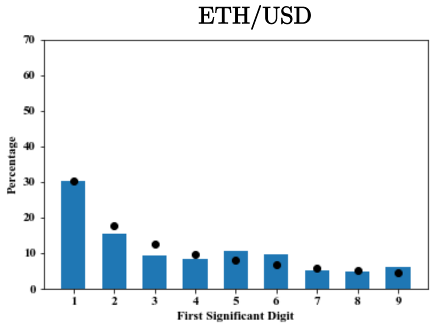


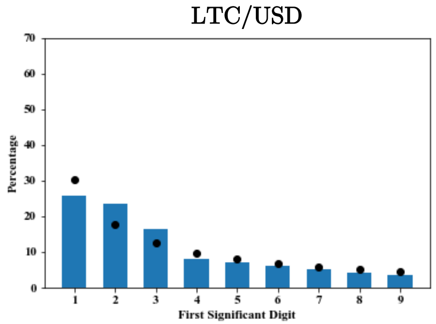


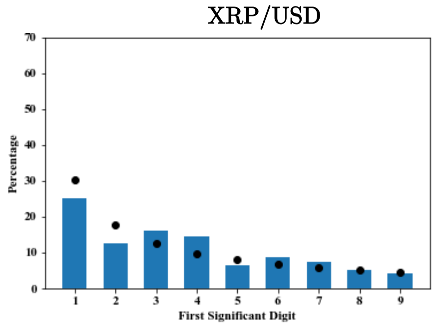
Panel UT: Unregulated Tier-1 exchanges

UT1

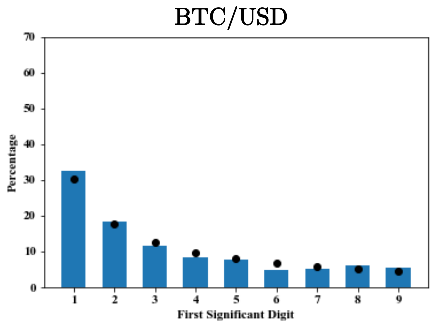


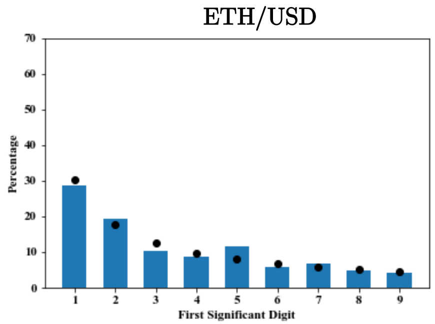


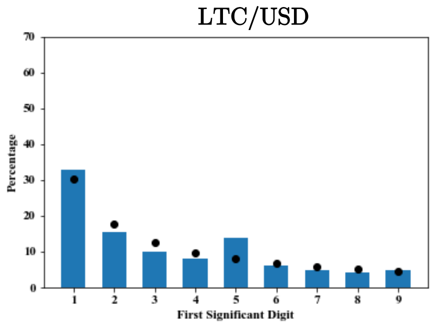


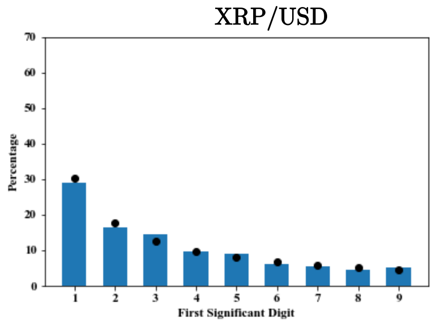


UT2

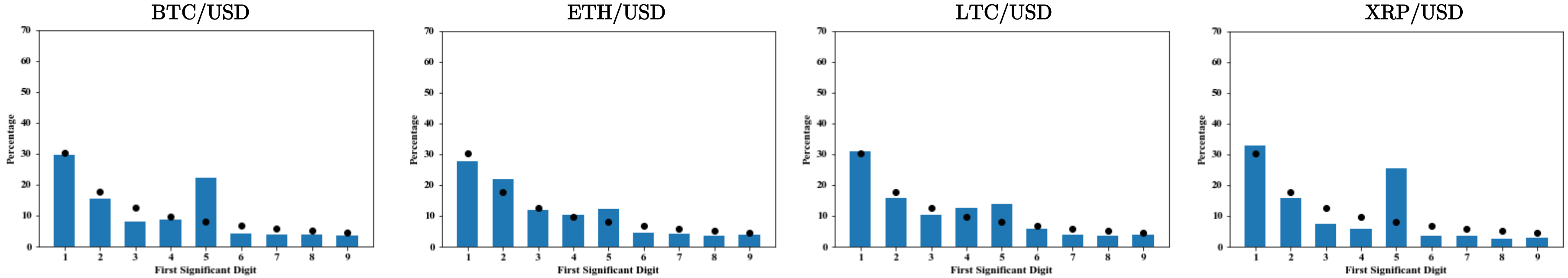




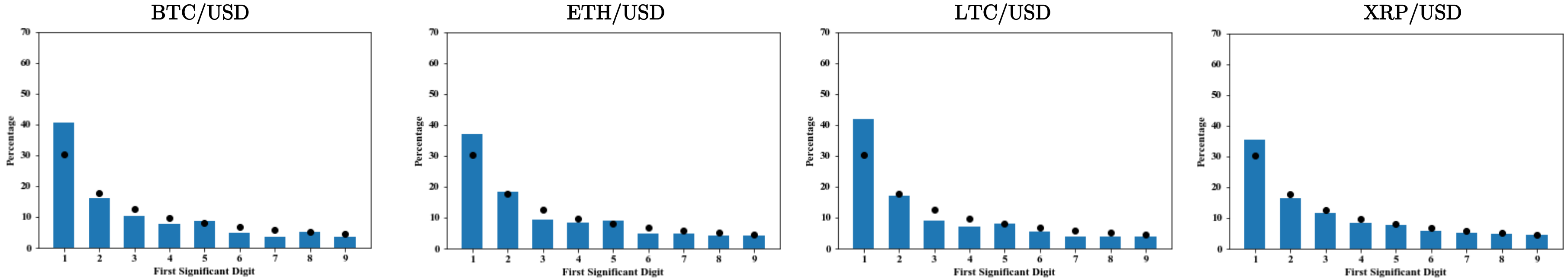




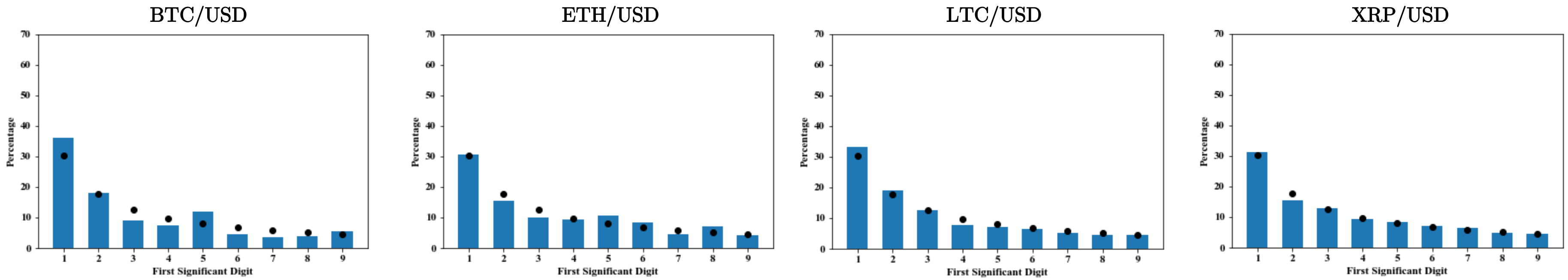
UT3


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

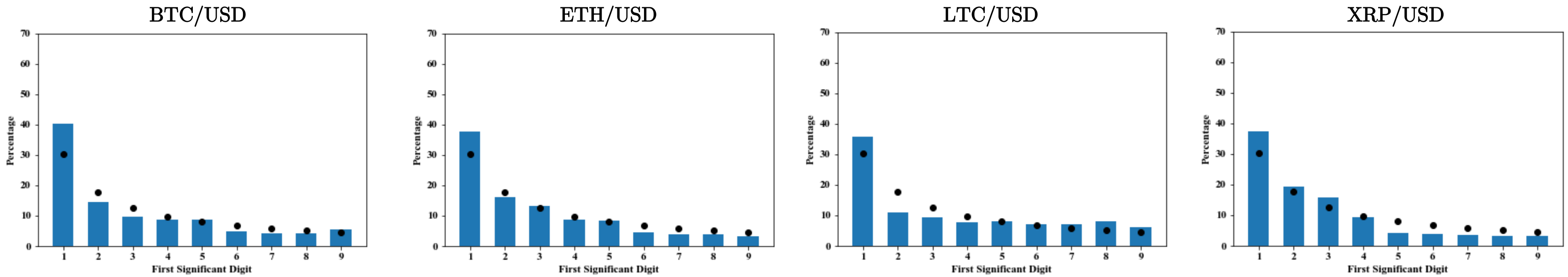
UT4


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

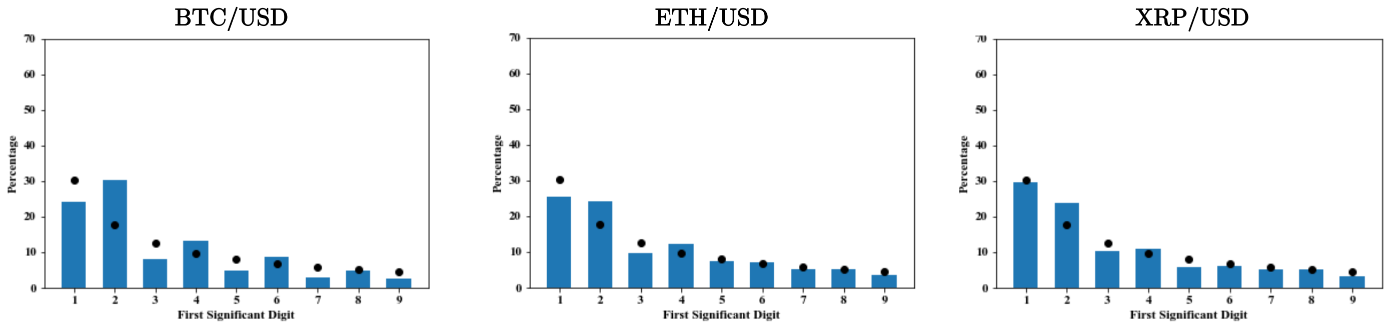
UT5


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

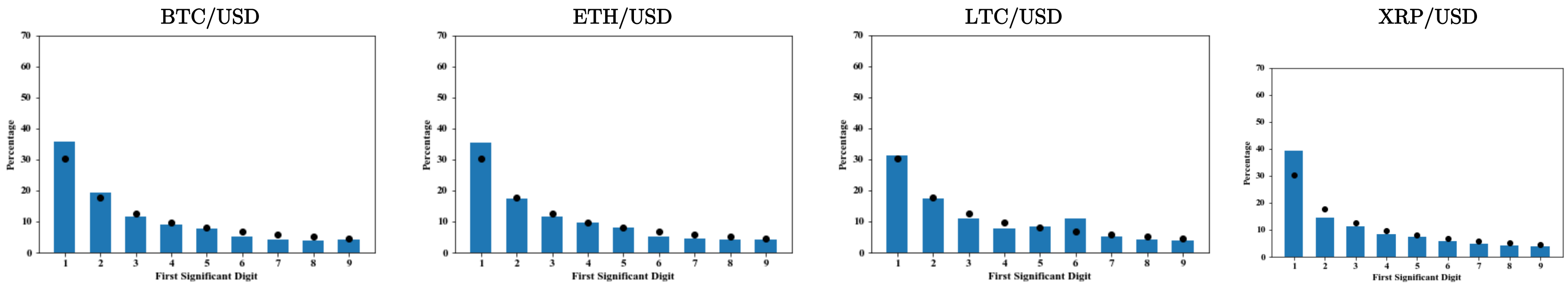
UT6


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

UT7

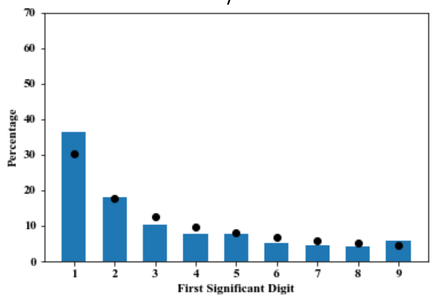

BTC/USD
ETH/USD
XRP/USD
Percentage
First Significant Digit

UT8

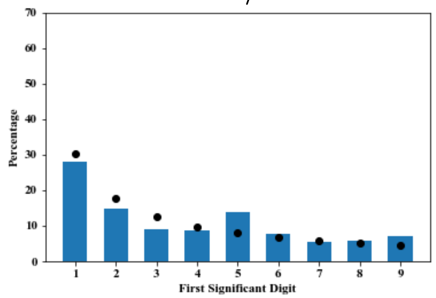

BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

UT9

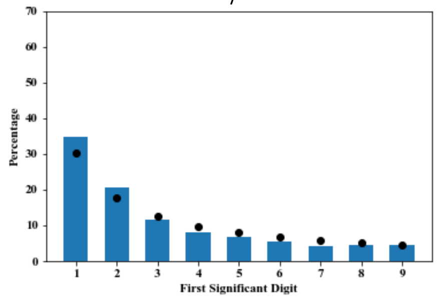
BTC/USD



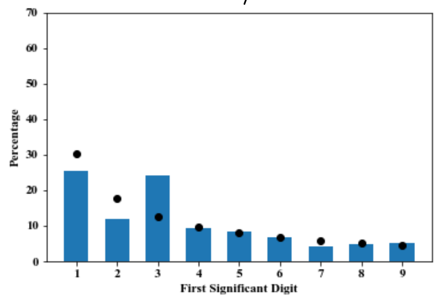
ETH/USD



LTC/USD

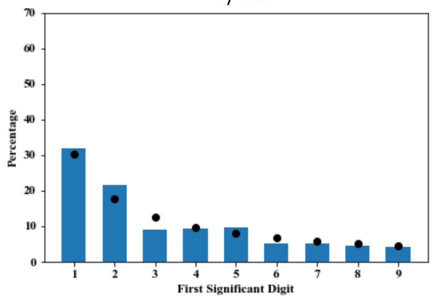


XRP/USD

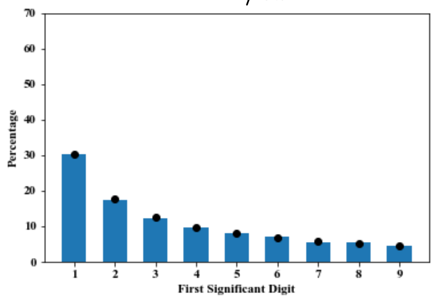


UT10

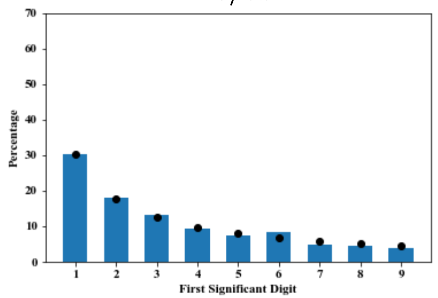
BTC/USD



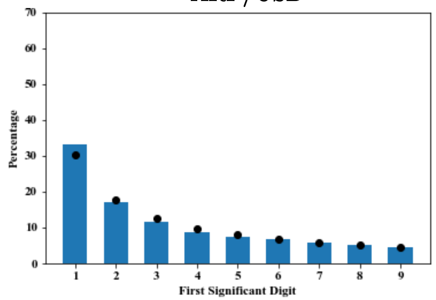
ETH/USD



LTC/USD



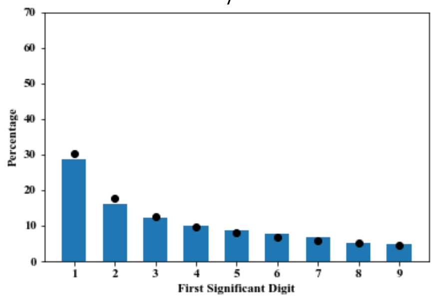
XRP/USD



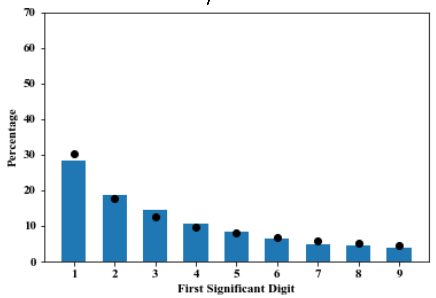
Panel U: Unregulated Tier-2 exchanges

U1

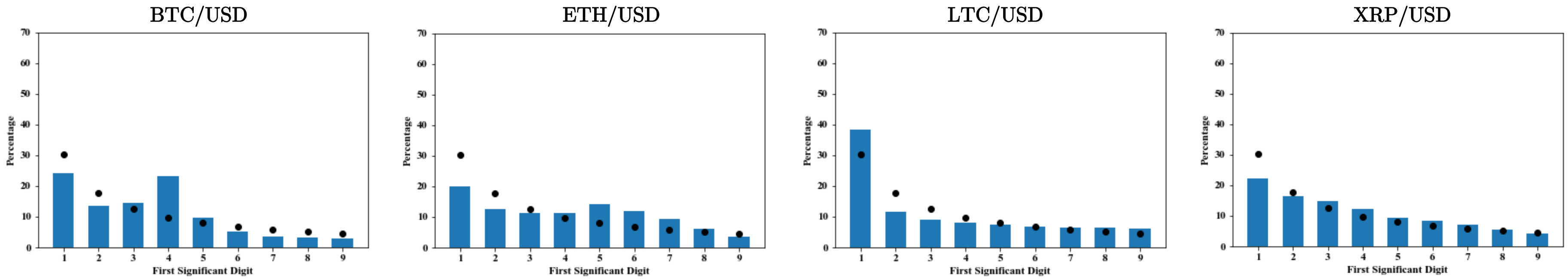
BTC/USD



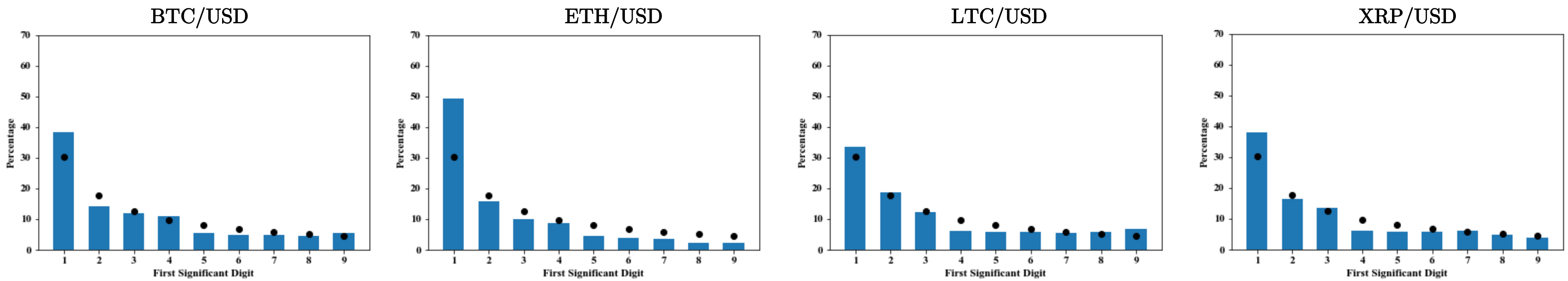
ETH/USD



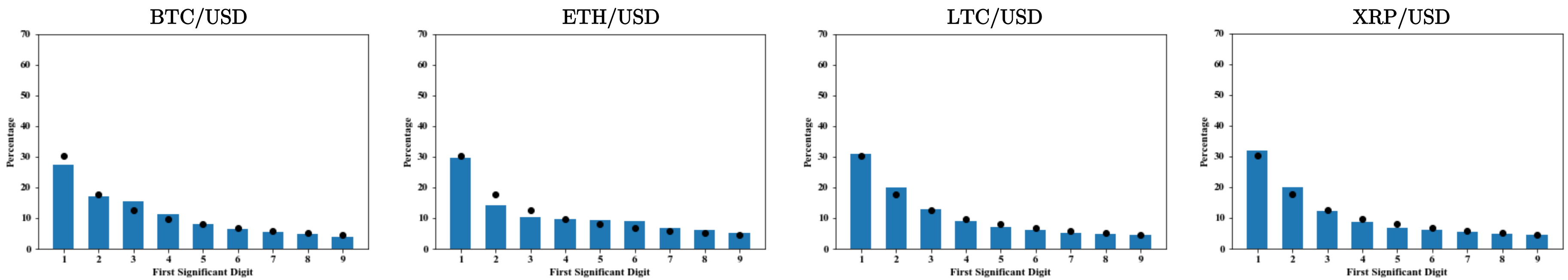
U2


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

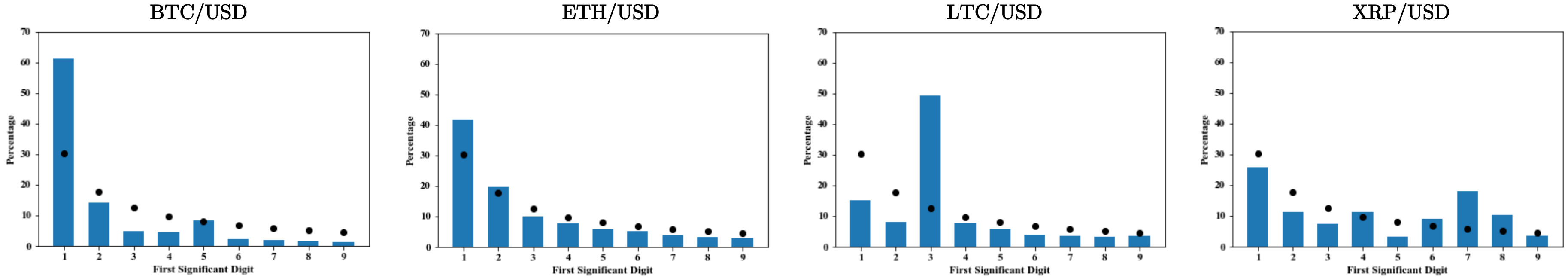
U3


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

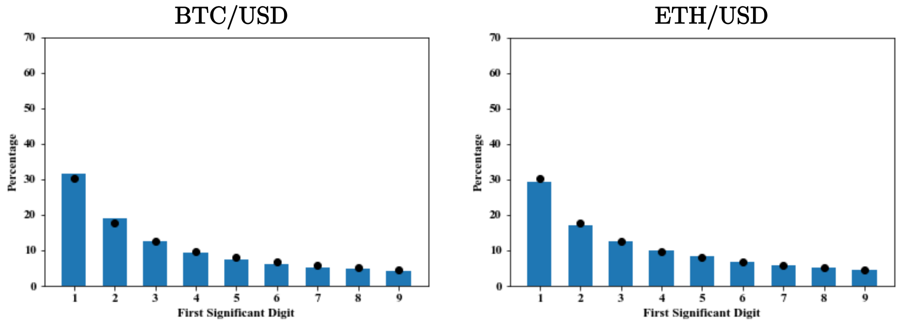
U4


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

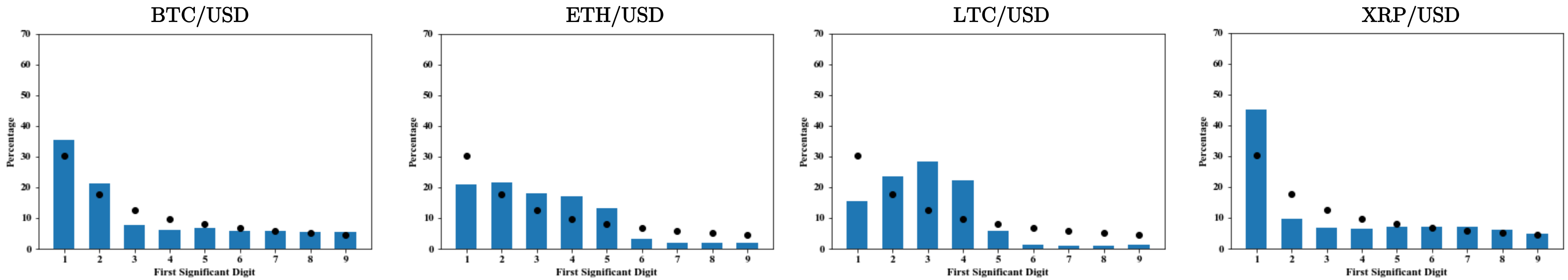
U5


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

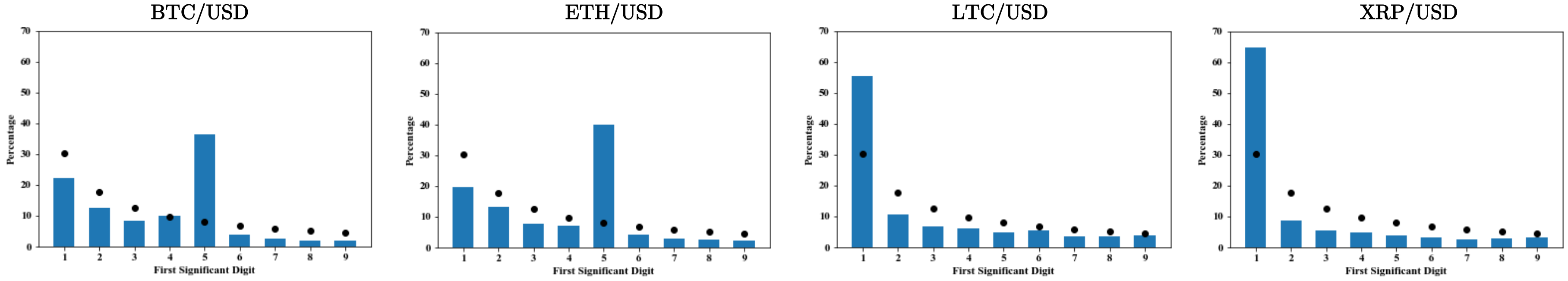
U6


BTC/USD
ETH/USD
Percentage
First Significant Digit

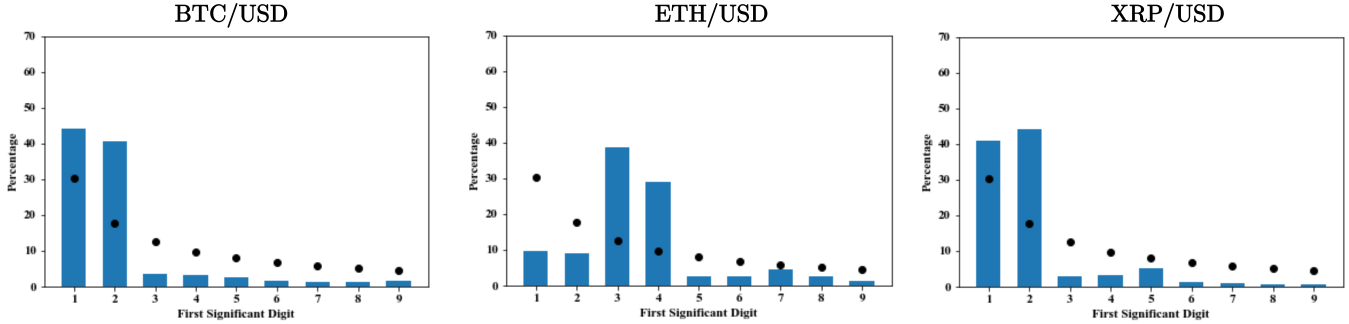
U7


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

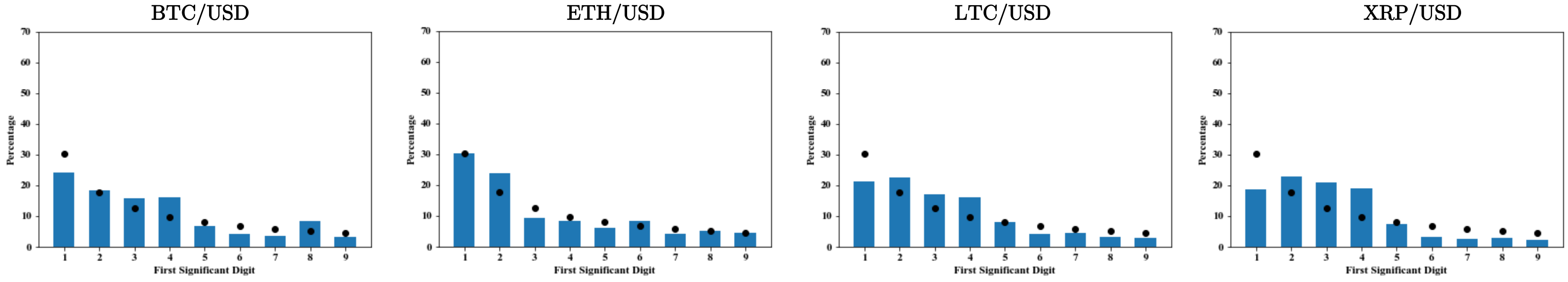
U8


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

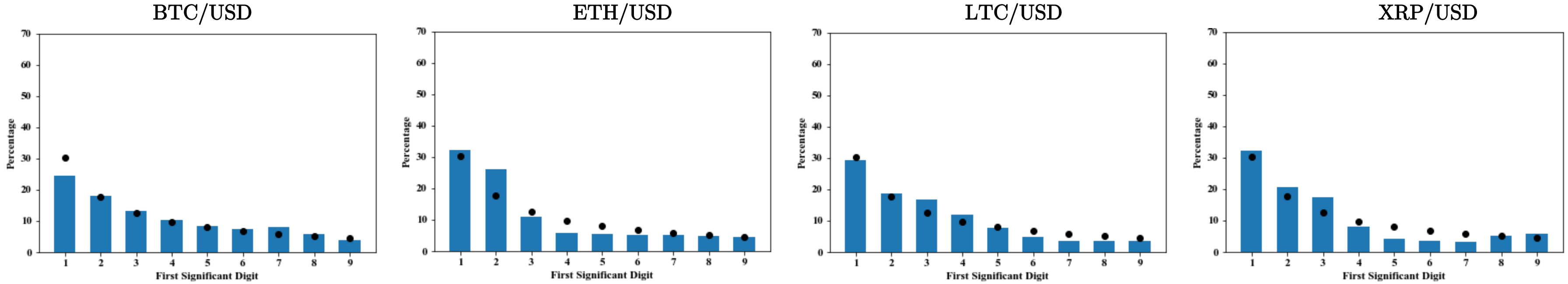
U9


BTC/USD
ETH/USD
XRP/USD
Percentage
First Significant Digit

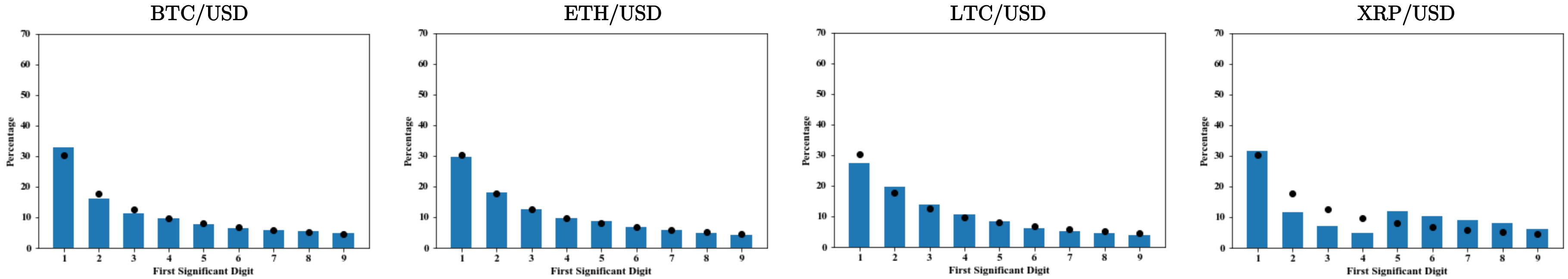
U10


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

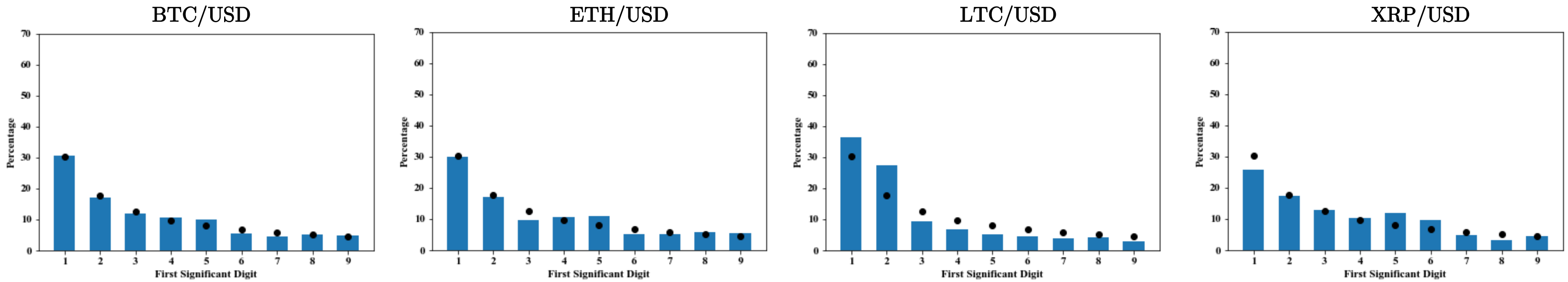
U11


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

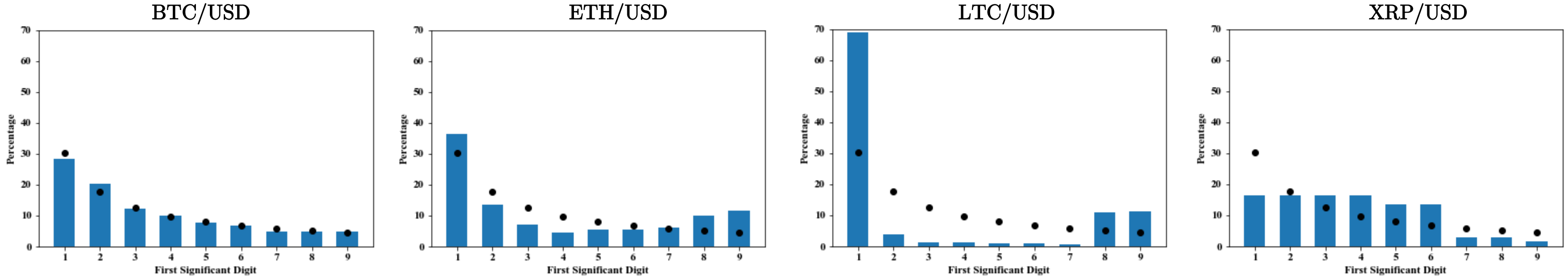
U12


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

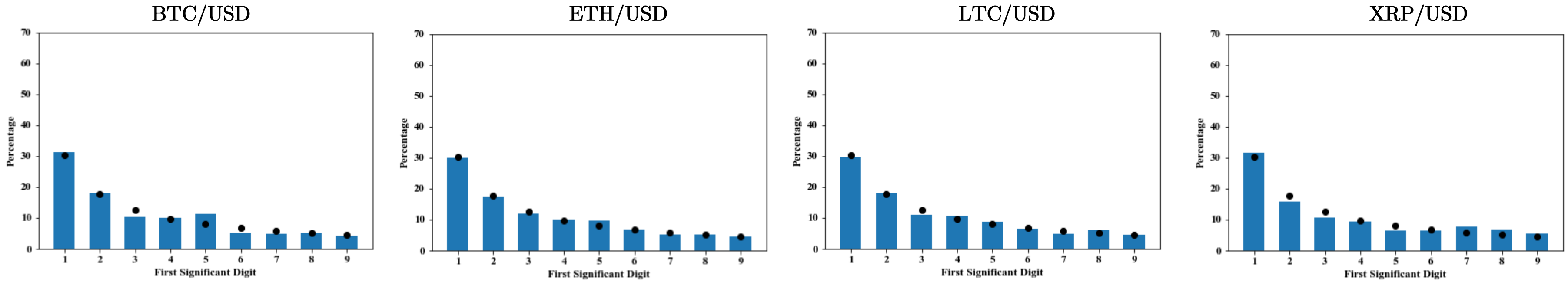
U13


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

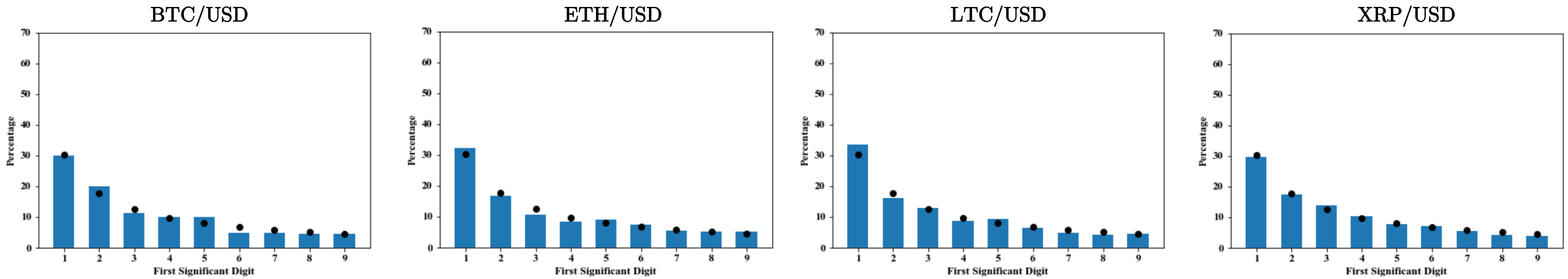
U14


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

U15


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

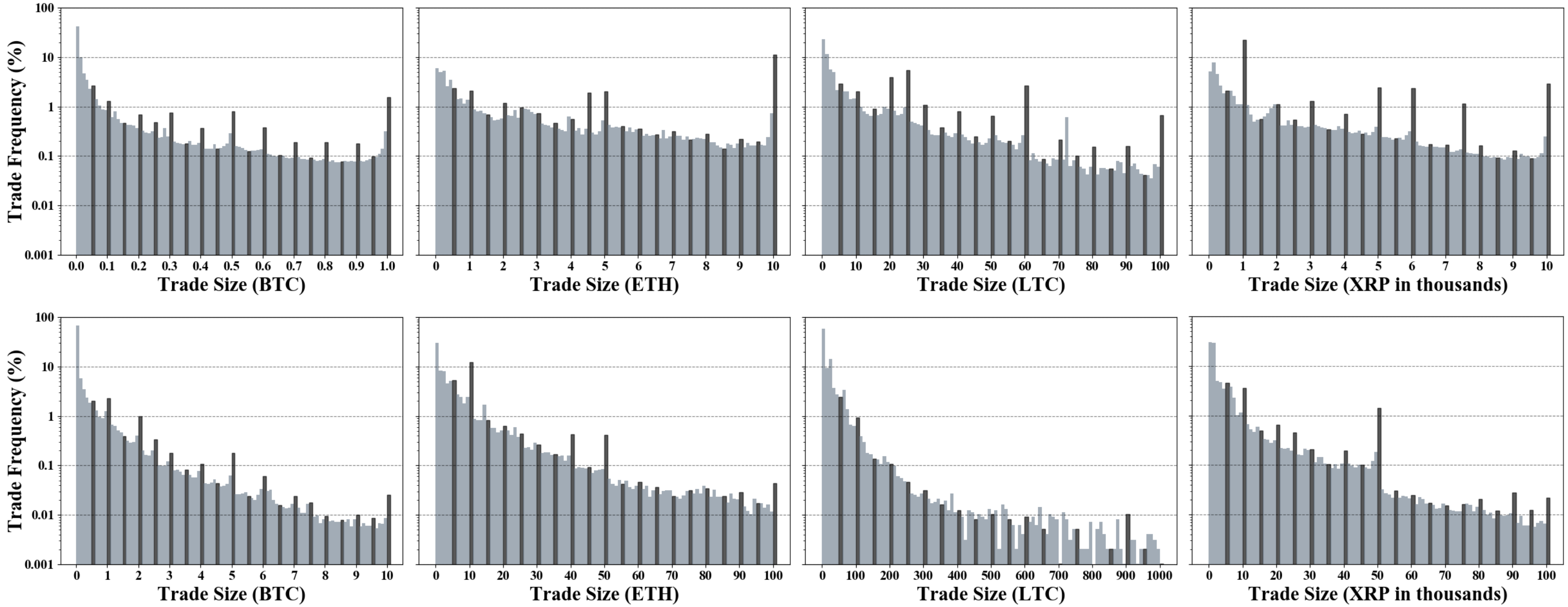
U16


BTC/USD
ETH/USD
LTC/USD
XRP/USD
Percentage
First Significant Digit

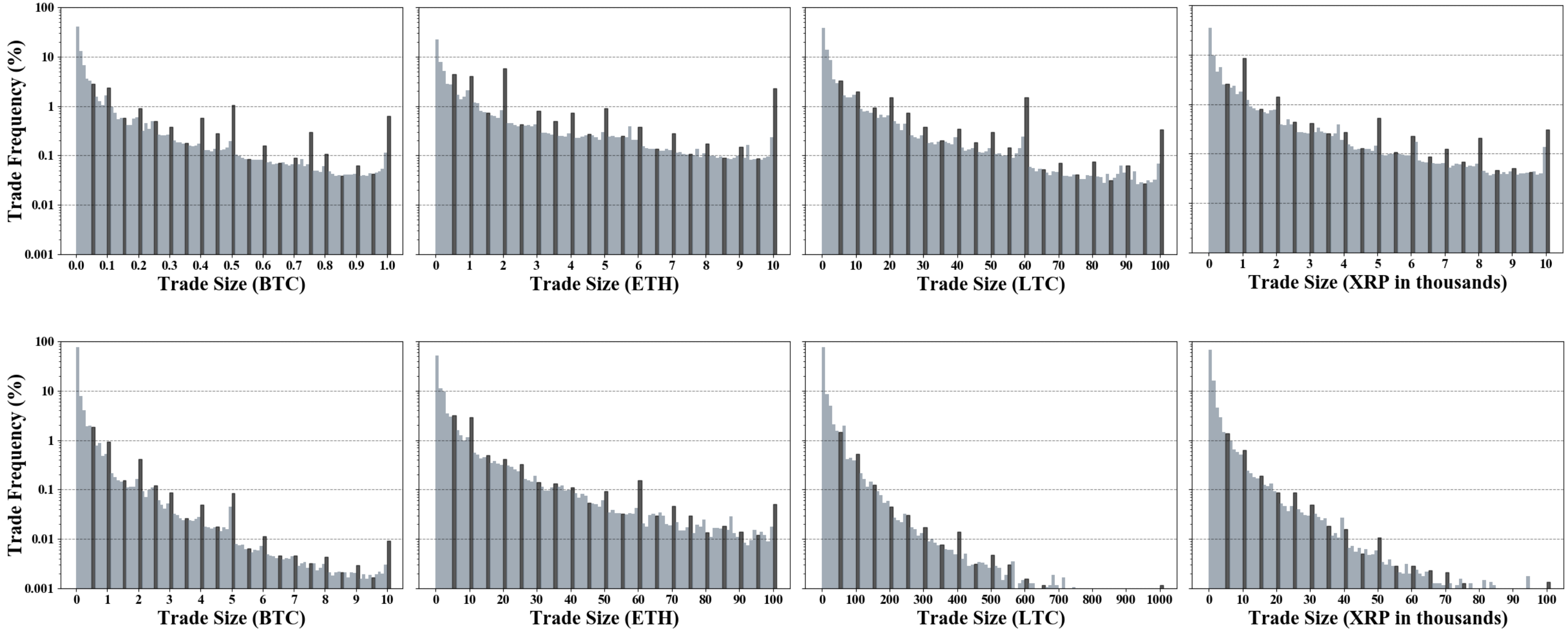
Appendix B Trade-size distribution and trade size clustering of crypto exchanges

Appendix B complements Figure 2 and Figure 3. Here displays the trade-size distribution and the size clustering figures of cryptocurrency exchanges not reported in Figure 2. . Distributions of four trading pairs are reported in form of histograms, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Four trading pairs are depicted over ranges of trade size, 0-1BTC, 0-10BTC, 0-10ETH, 0-100ETH, 0-100LTC, 0-1000LTC, 0-10,000XPR and 0-100,000XPR. Every 5th and 10th bins are highlighted to illustrate the clustering effect around rounded trade sizes. Y-axis represents the frequency of trades on a log scale. Panel R, Panel UT, and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

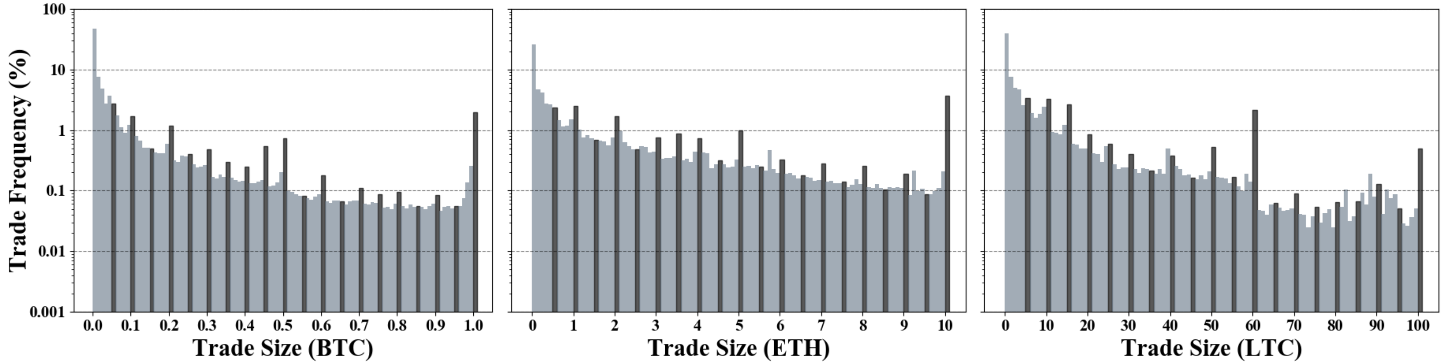
Panel R: Regulated exchanges
R1

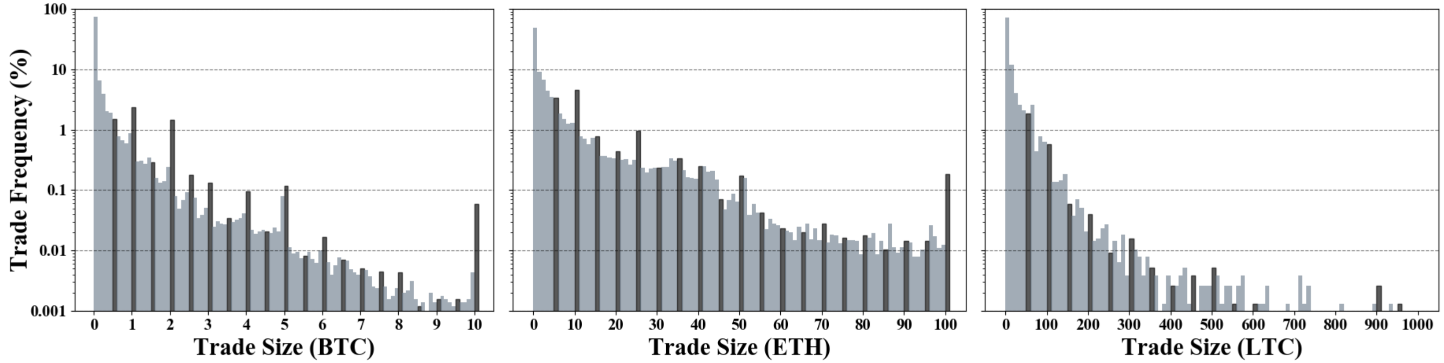


R2


100
10
1
0.1
0.01
0.001
Trade Frequency (%)
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)
100
10
1
0.1
0.01
0.001
Trade Frequency (%)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (BTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (ETH)
0 100 200 300 400 500 600 700 800 900 1000
Trade Size (LTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (XRP in thousands)

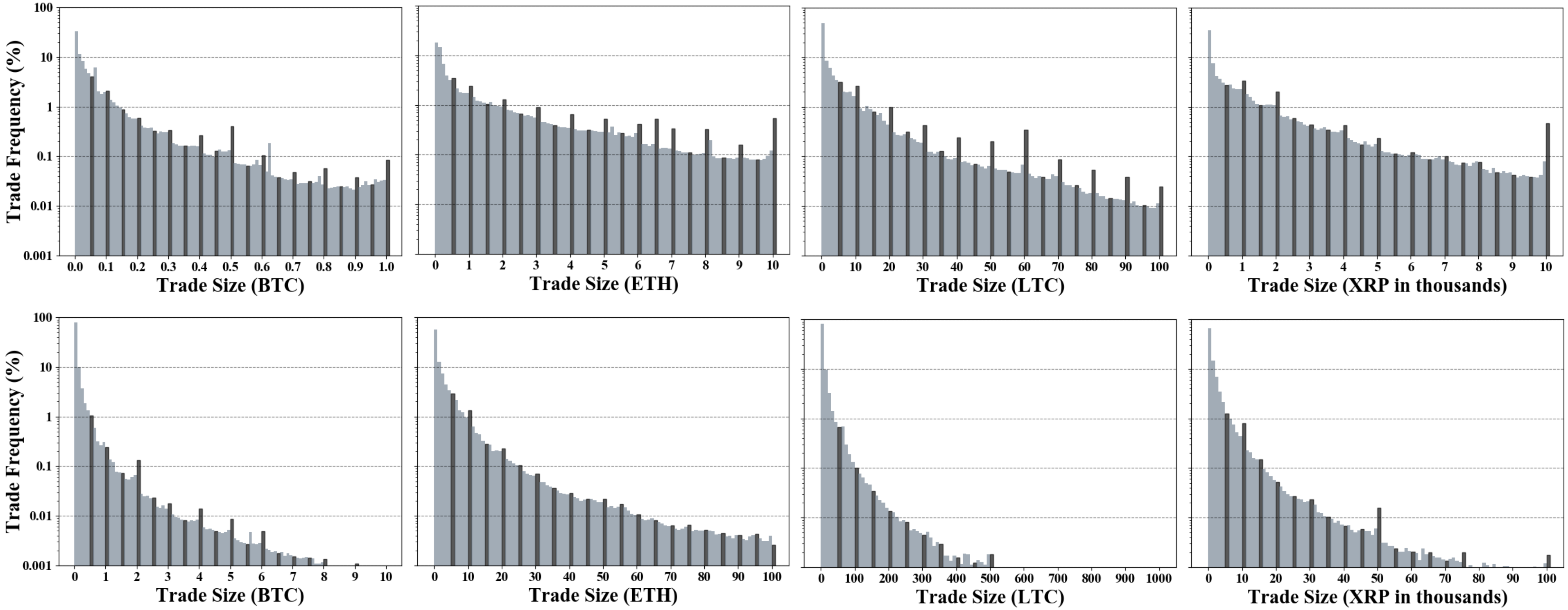
R3


100
10
1
0.1
0.01
0.001
Trade Frequency (%)
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)

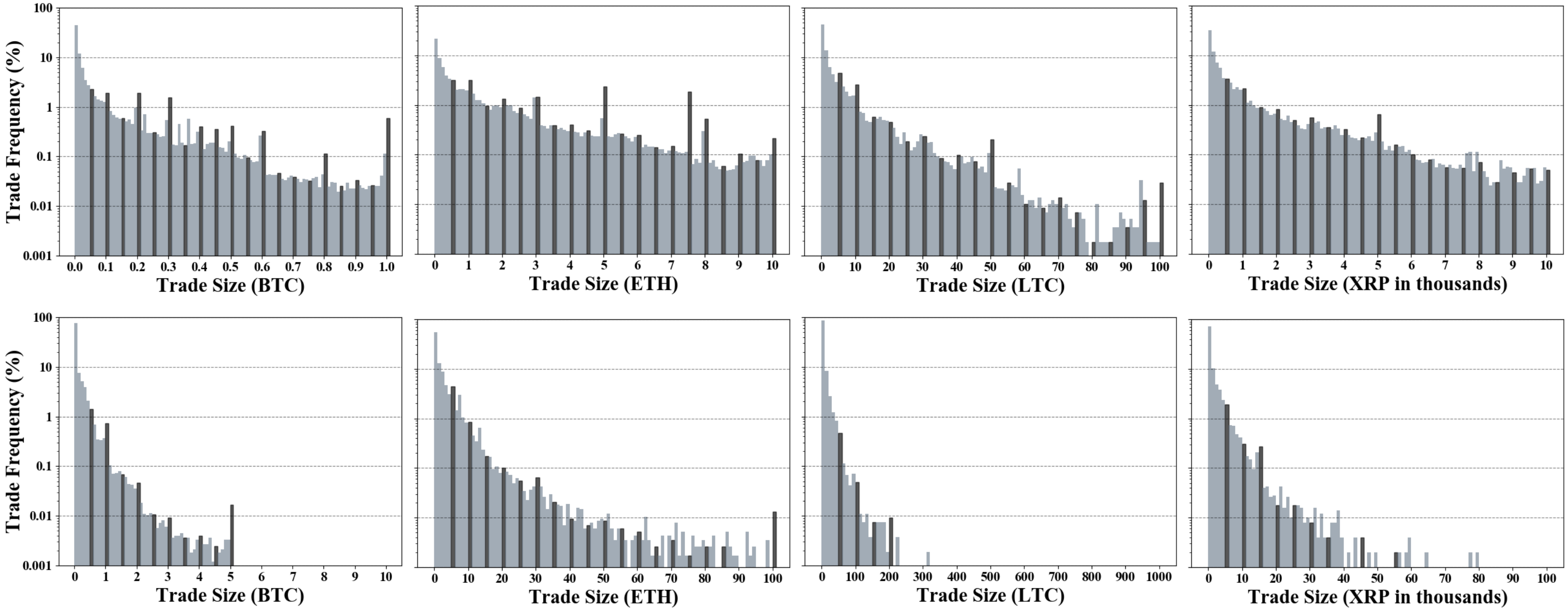


Panel UT: Unregulated Tier-1 exchanges

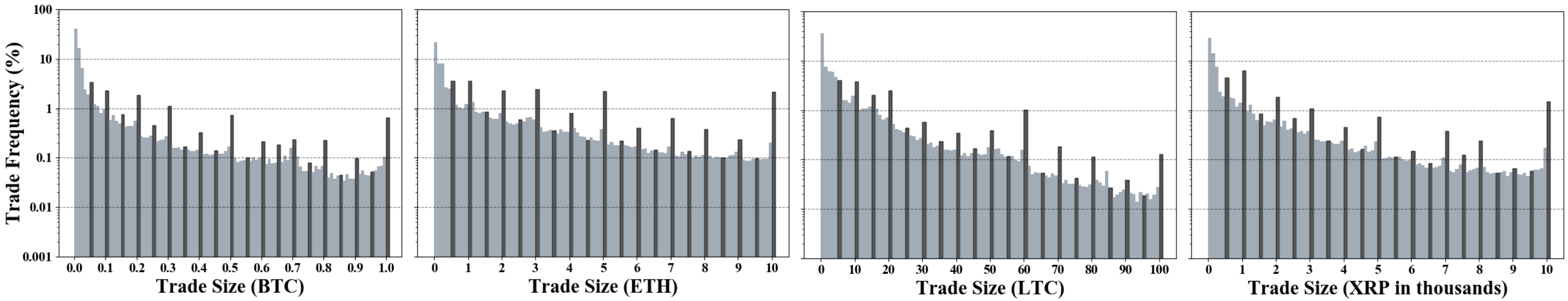
UT1

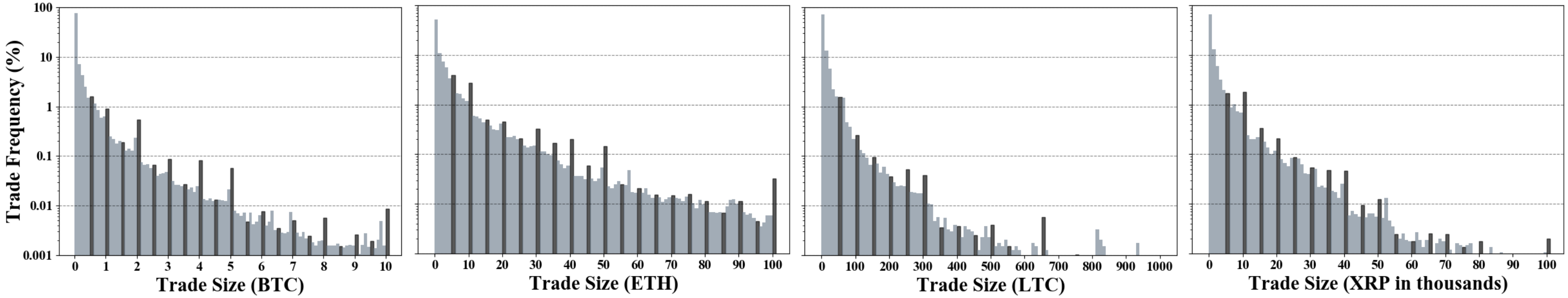


UT2

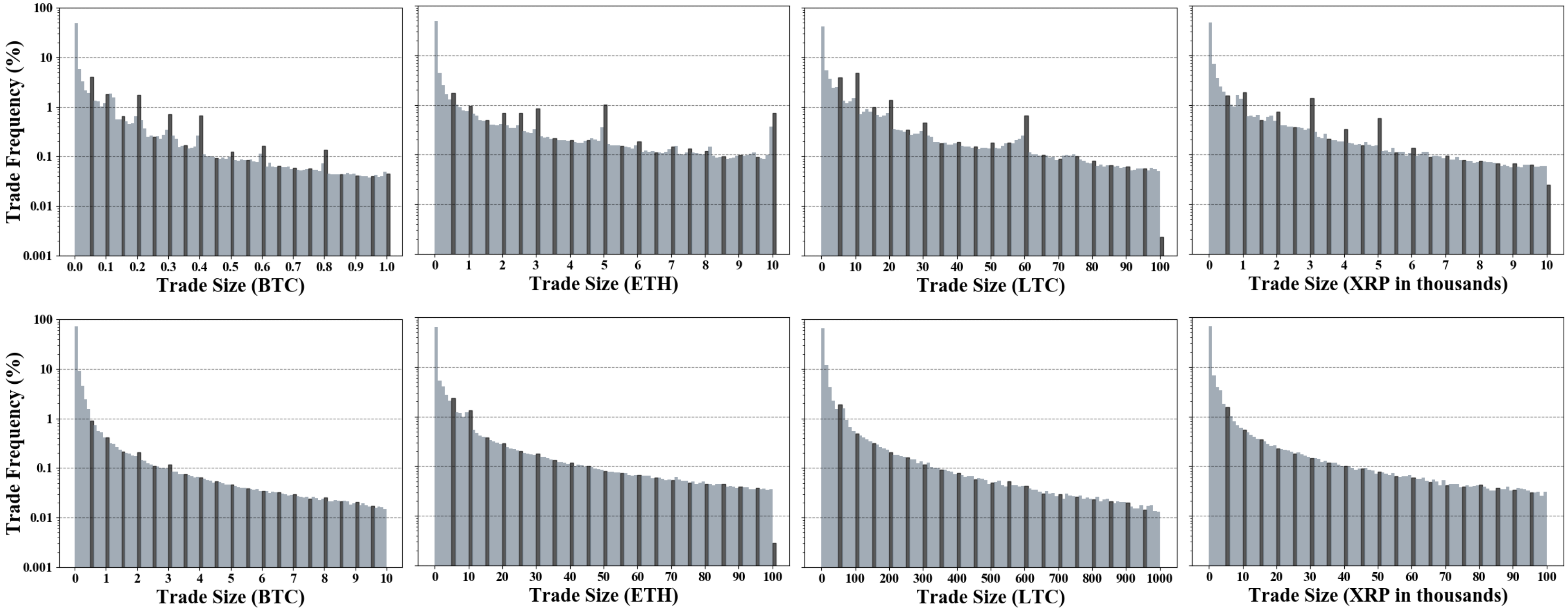

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

UT3

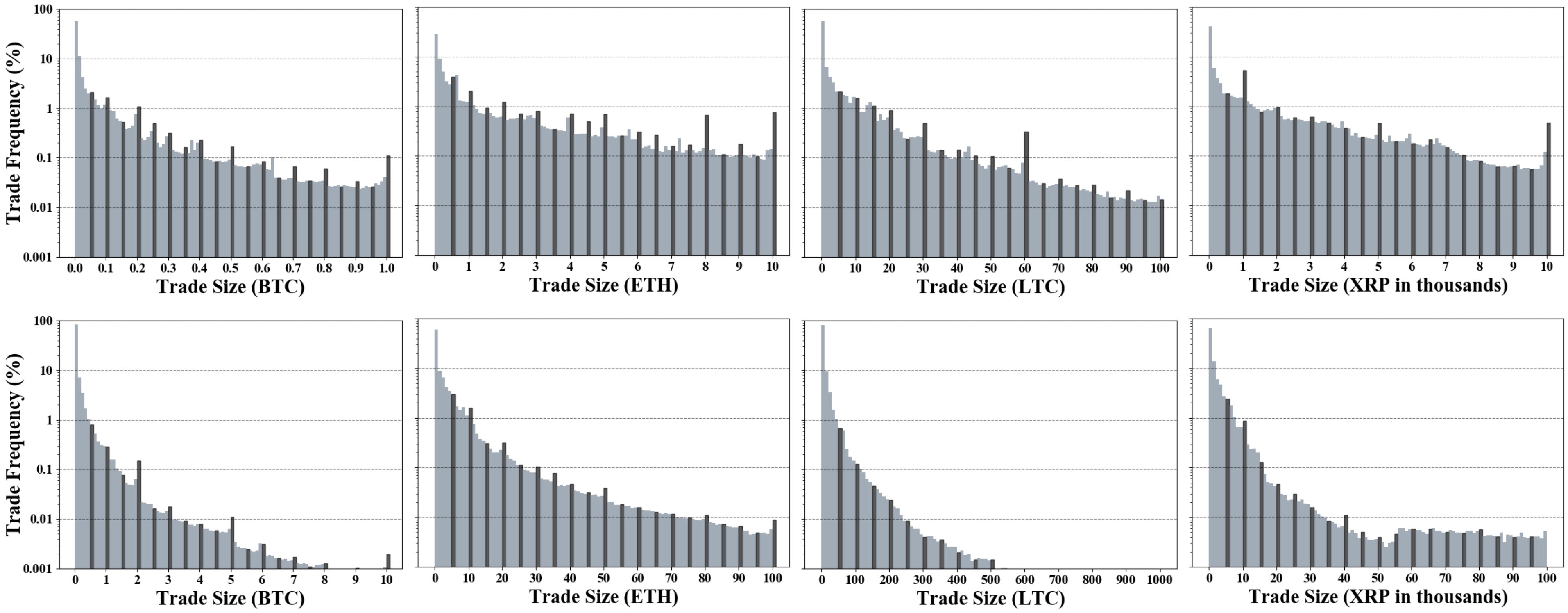

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)


Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

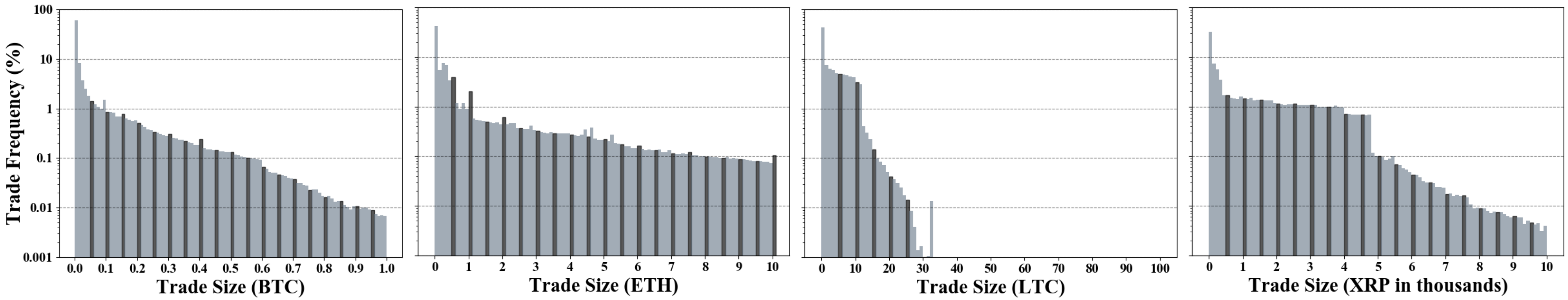
UT4

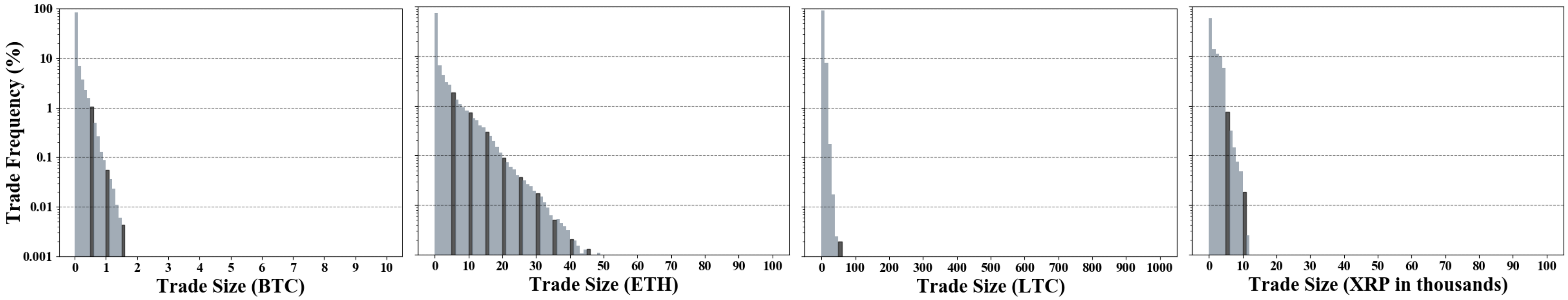

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

UT5

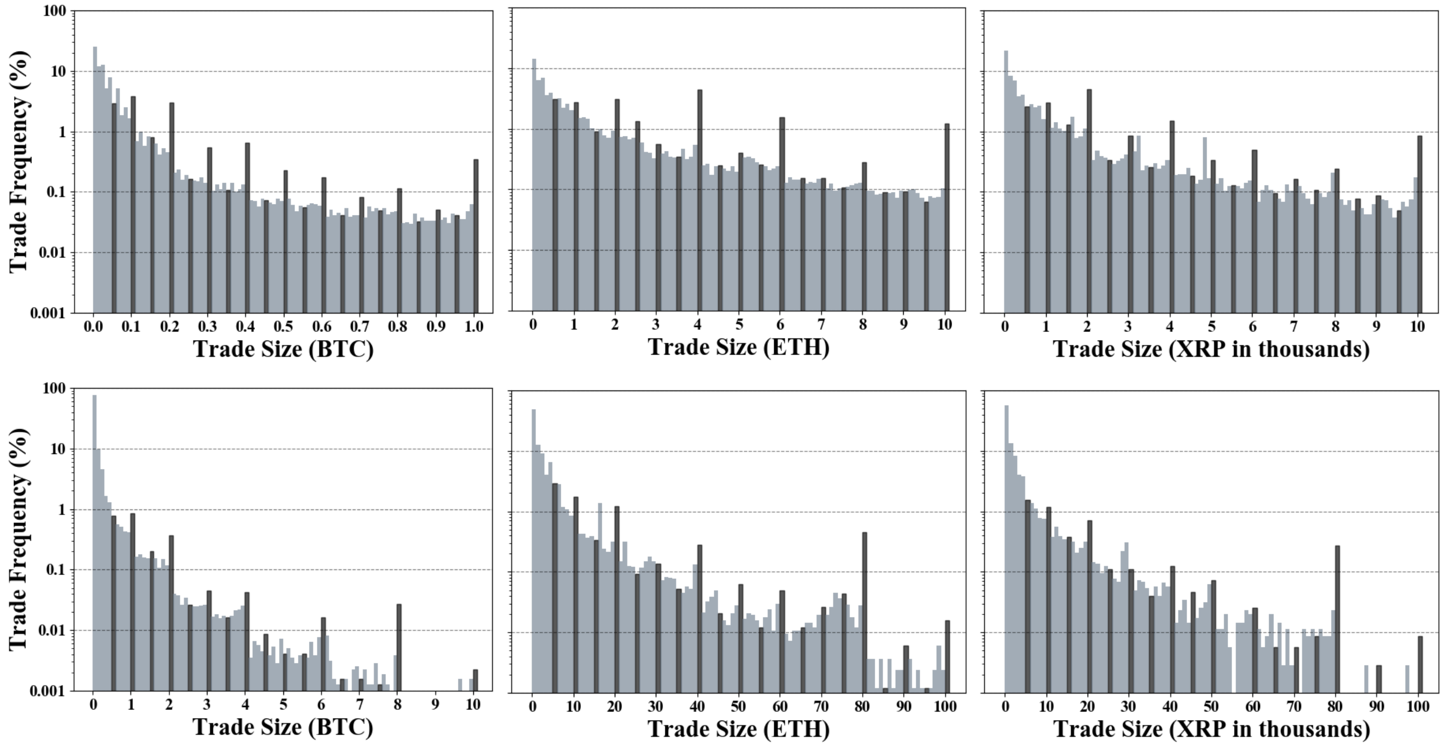

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

UT6

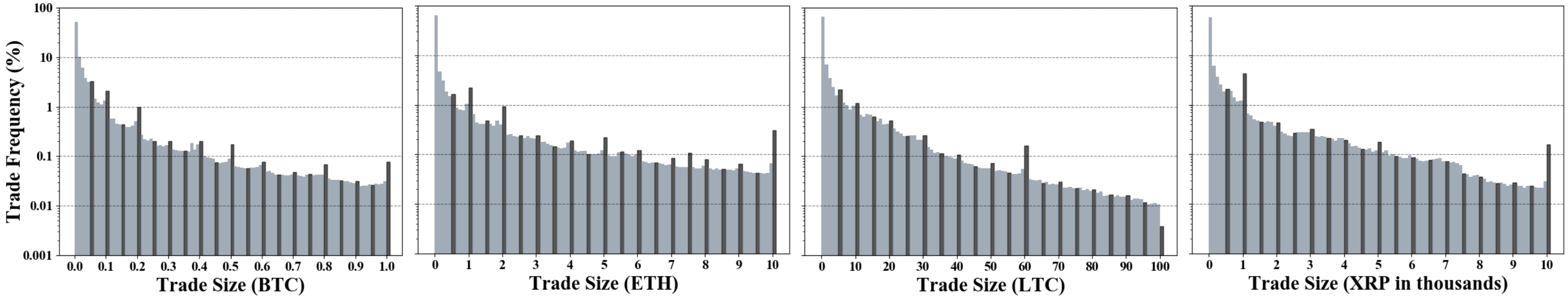

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

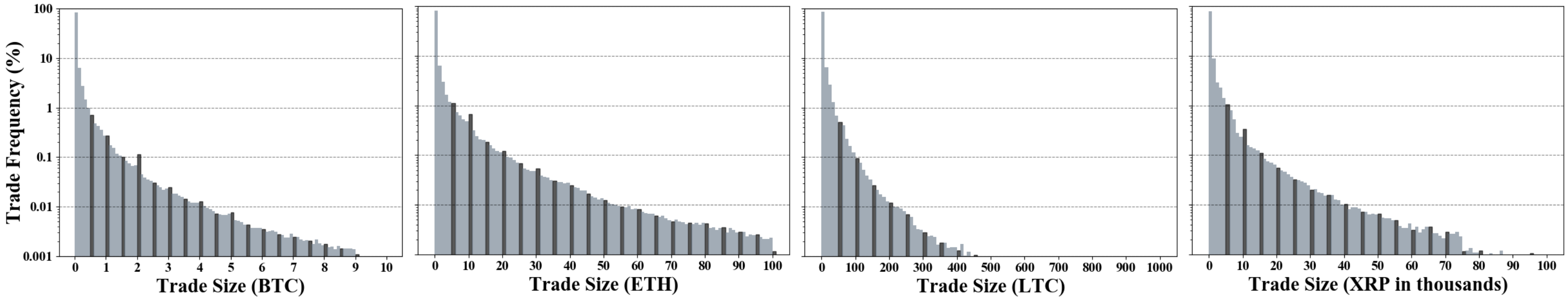

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

UT7

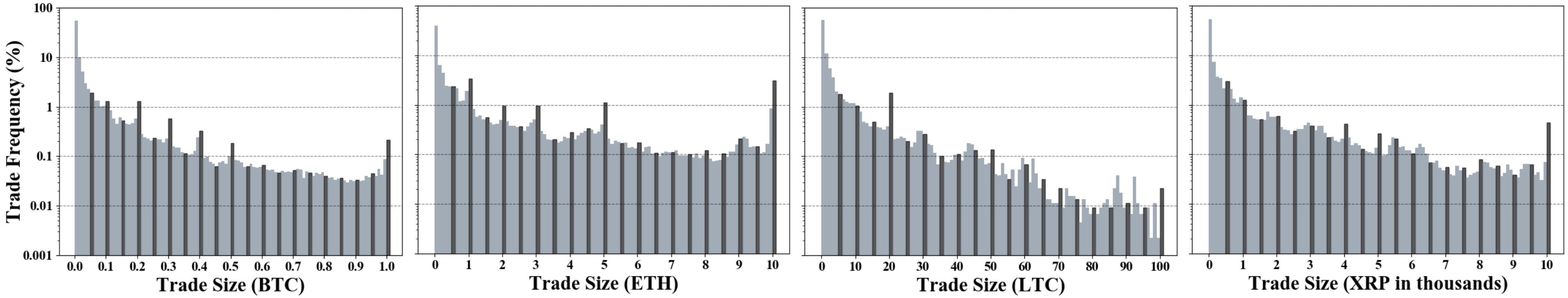

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (XRP in thousands)

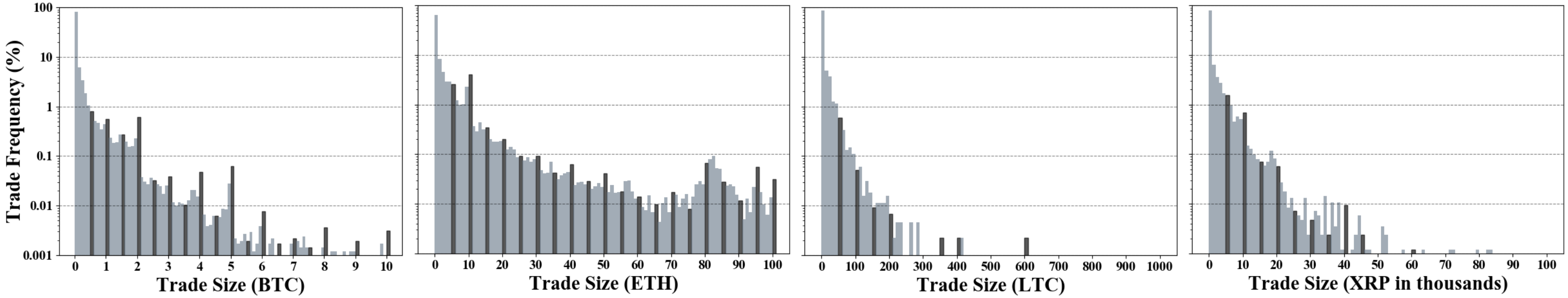
UT8


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)

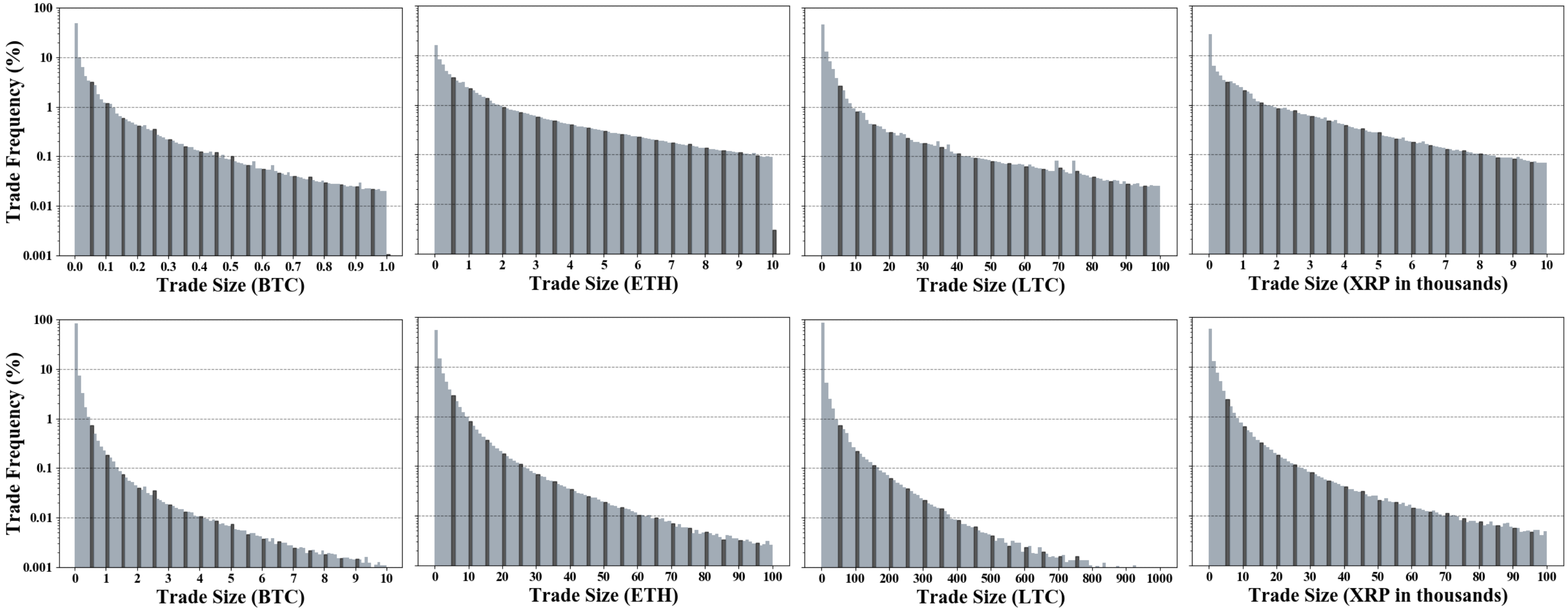

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0 1 2 3 4 5 6 7 8 9 10
Trade Size (BTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (ETH)
0 100 200 300 400 500 600 700 800 900 1000
Trade Size (LTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (XRP in thousands)

UT9


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)

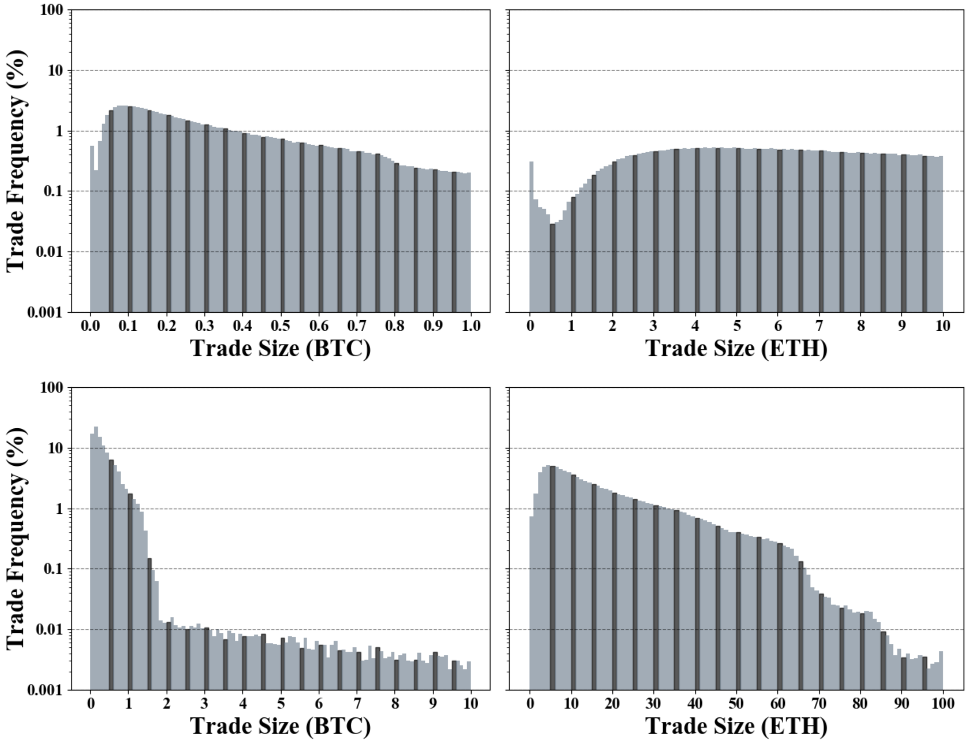

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

UT10

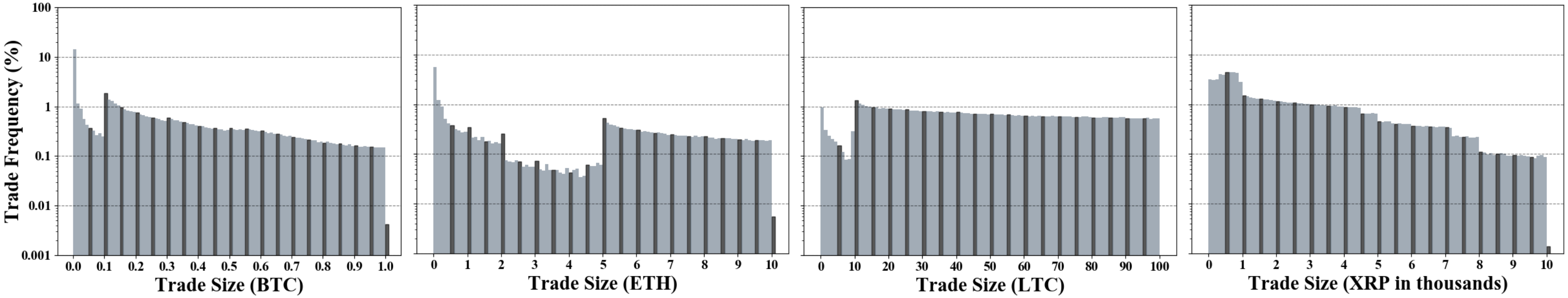

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

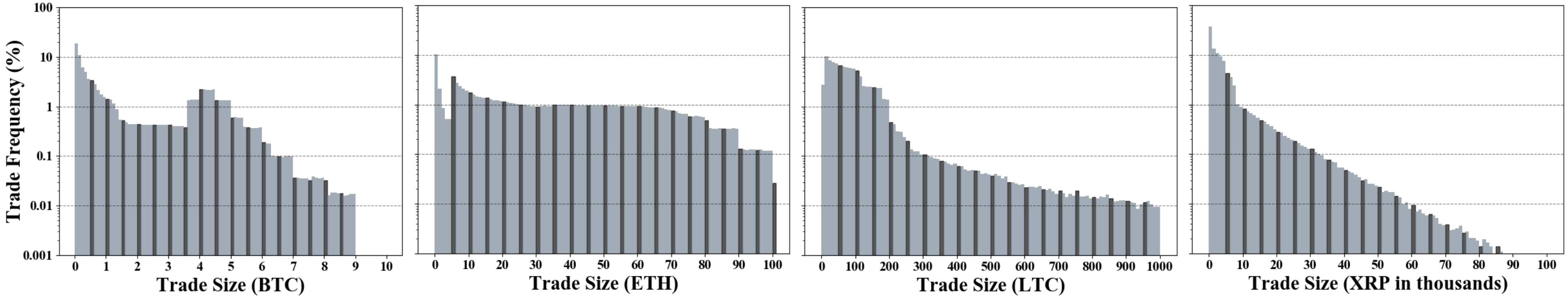
Panel U: Unregulated Tier-2 exchanges

U1

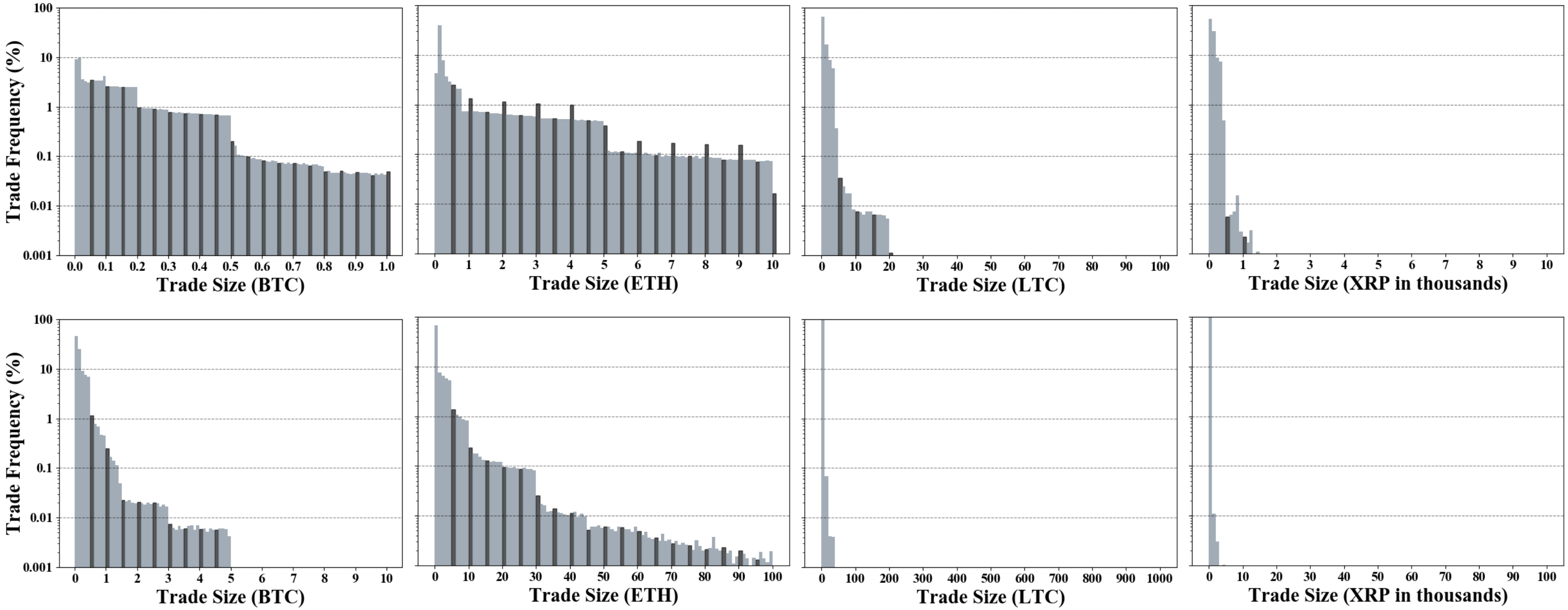


U2

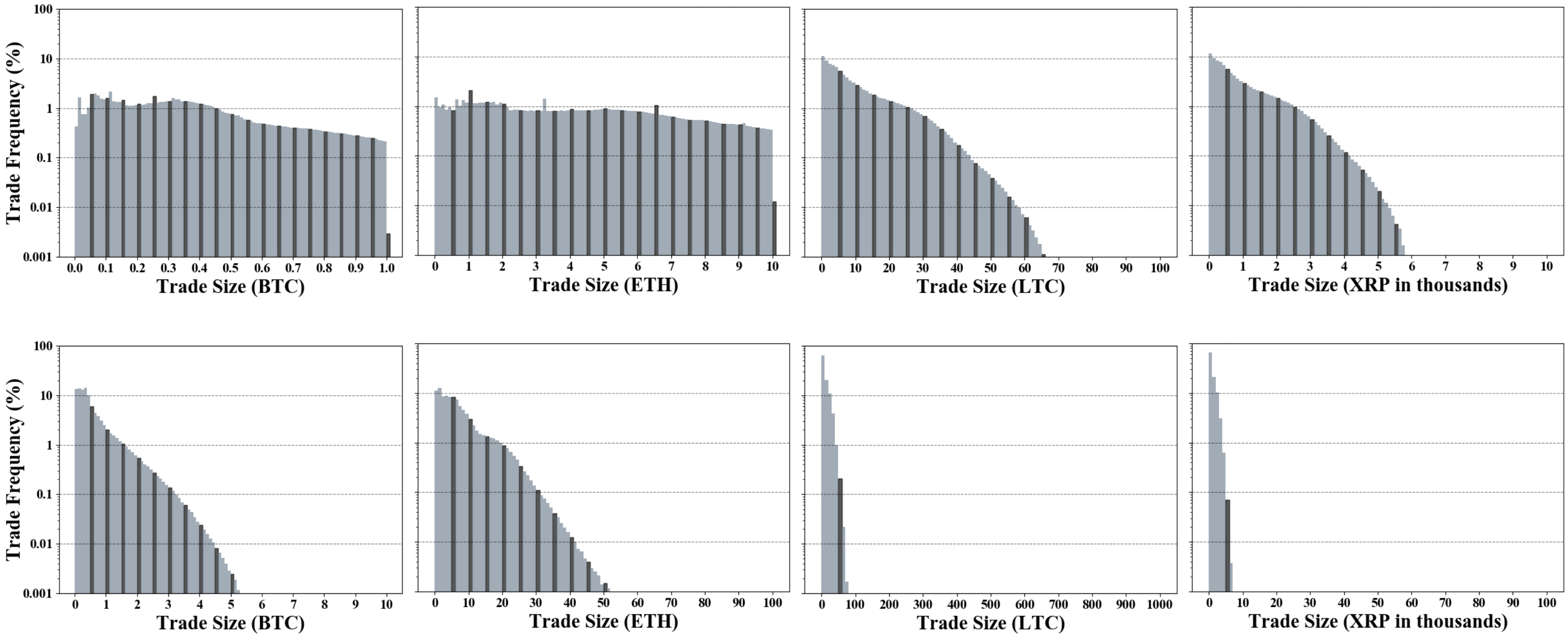



Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

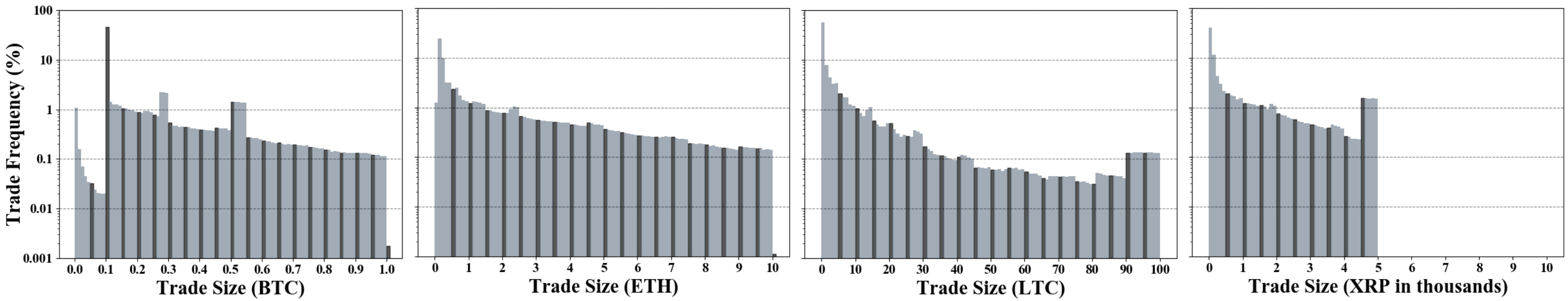
U3

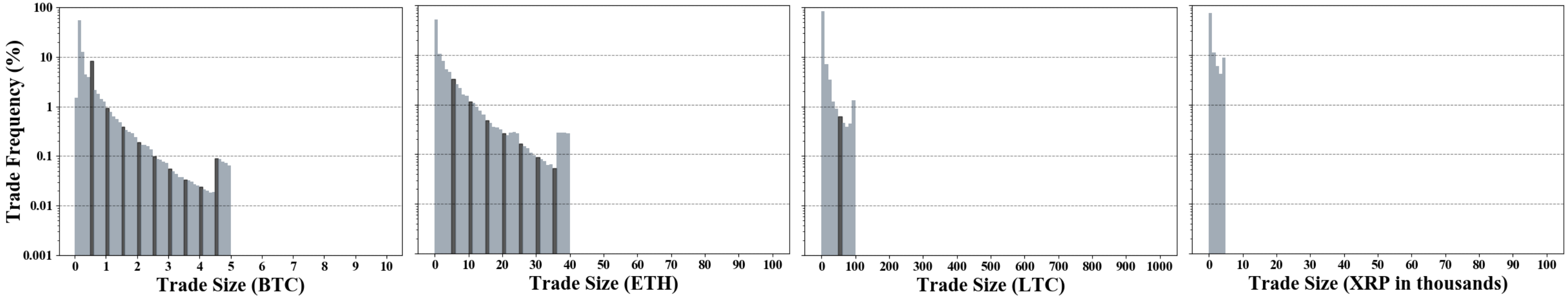

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U4


Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

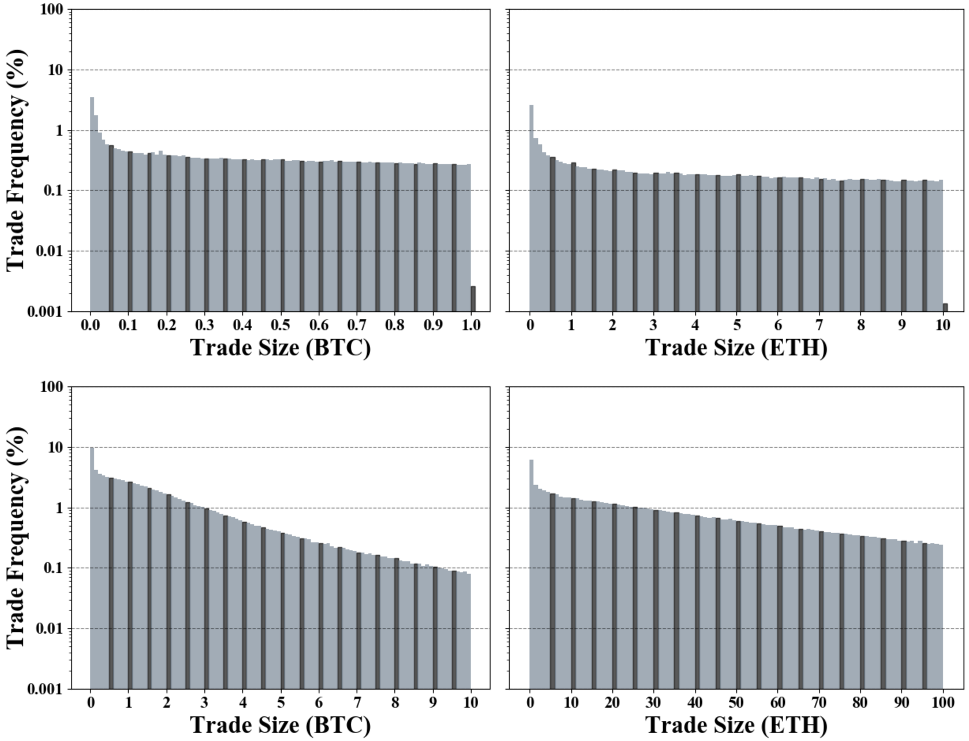
U5


Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)



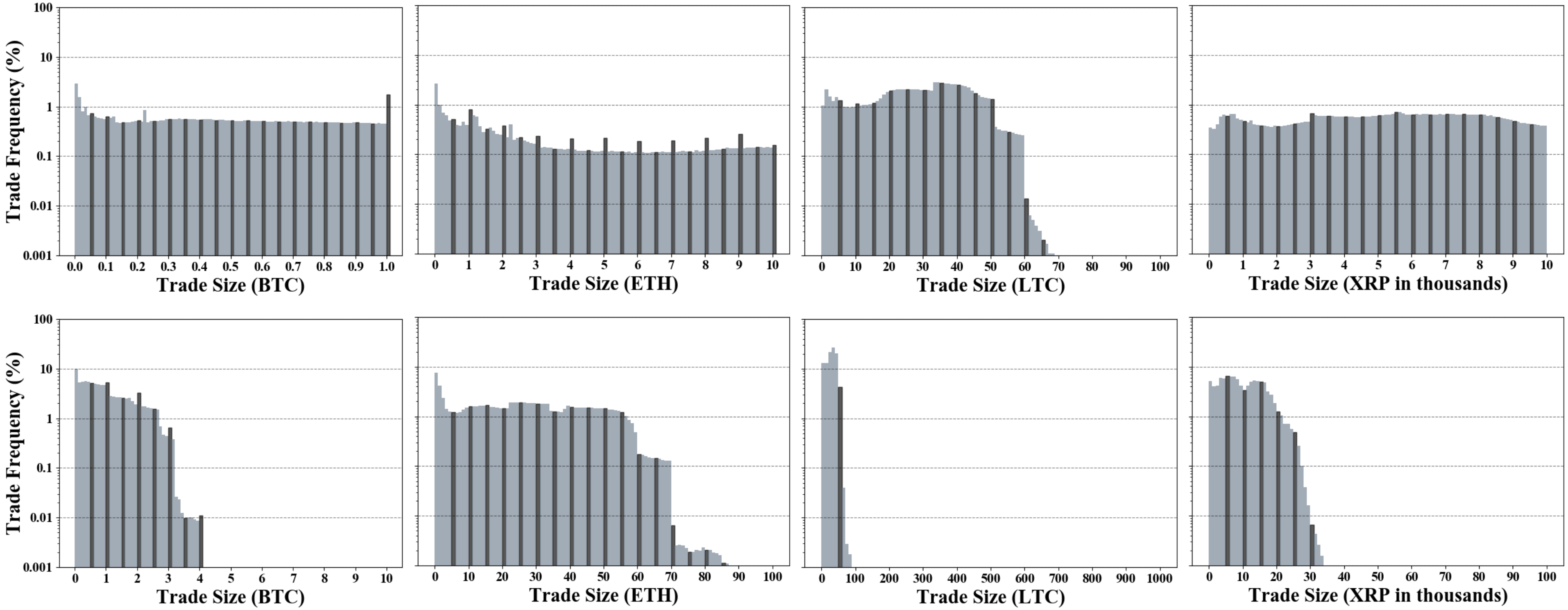
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U6

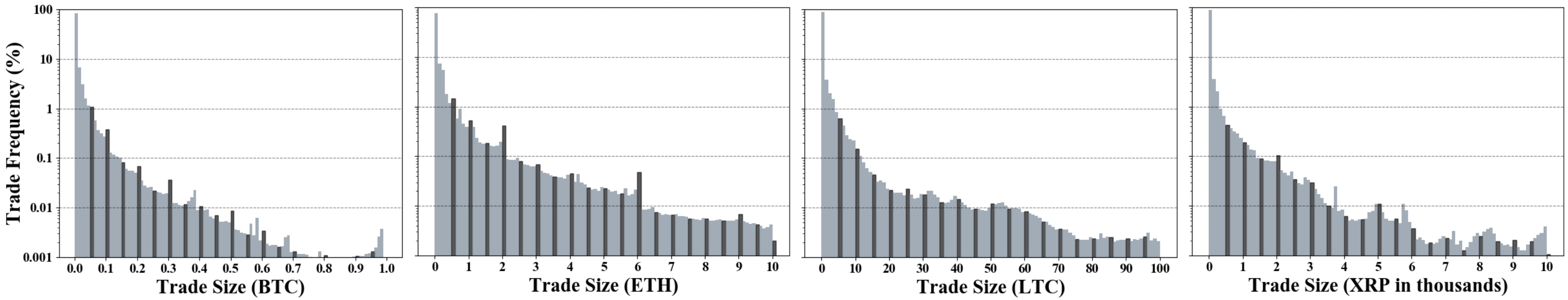


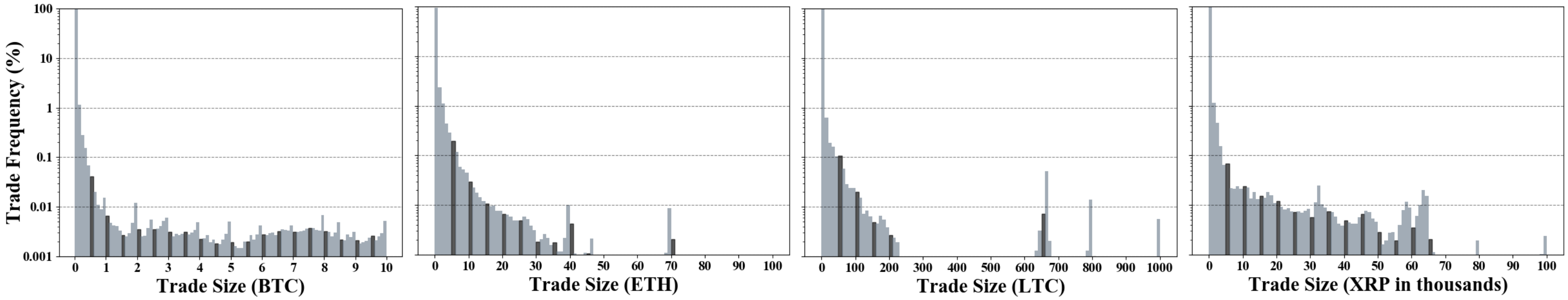
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)

U7

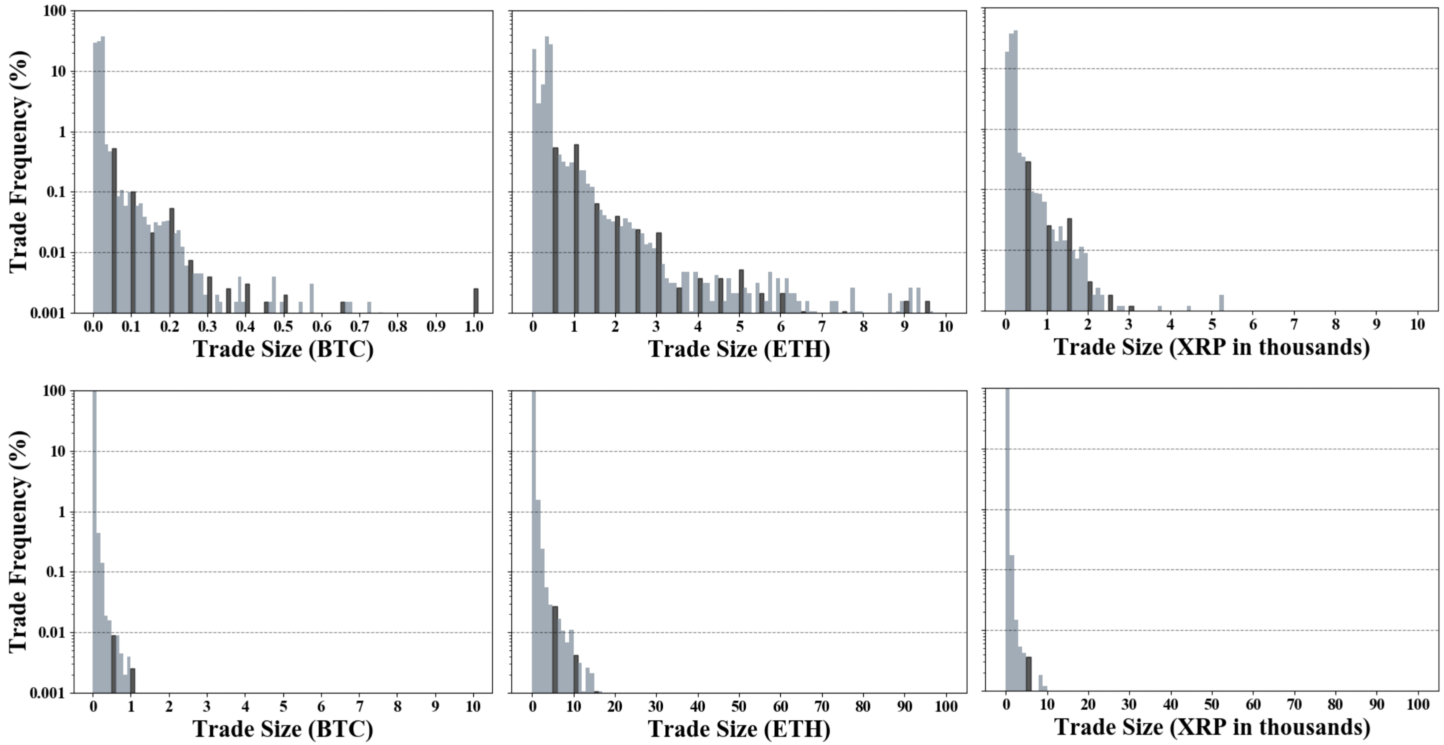

Trade Frequency (%)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U8

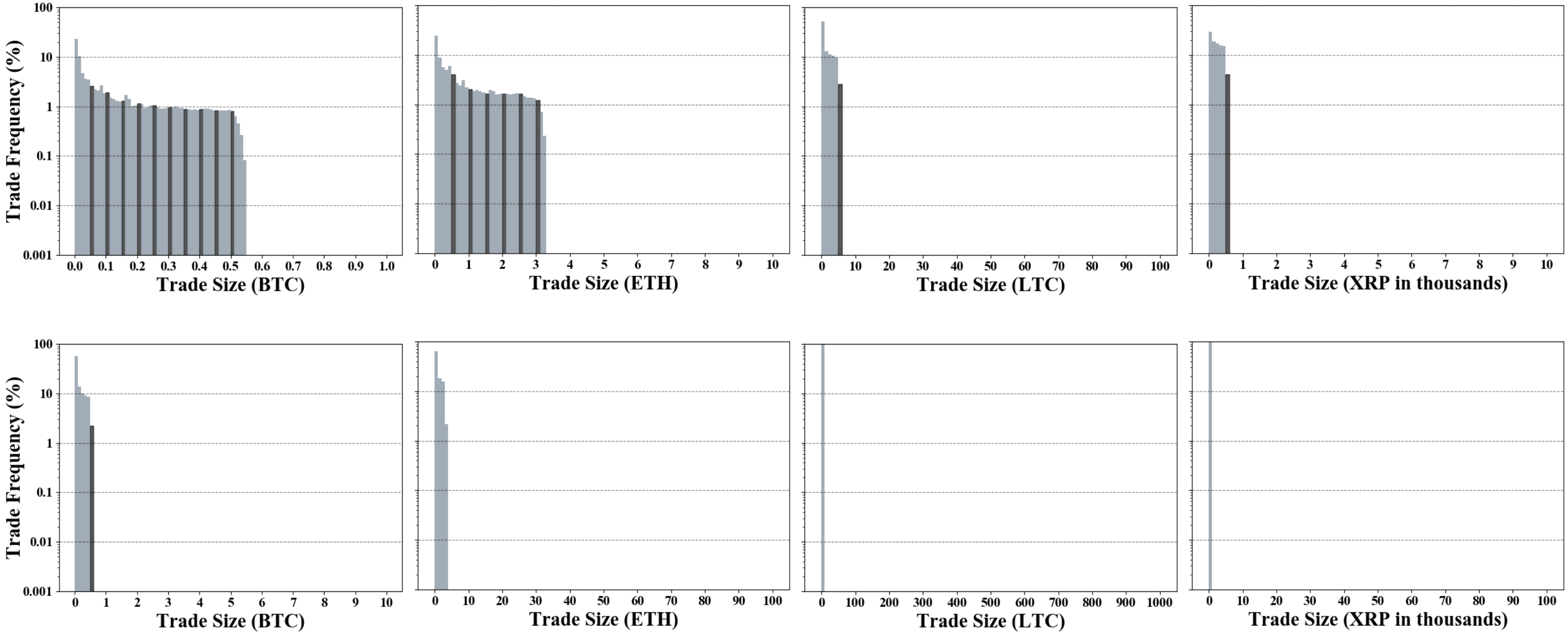

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U9

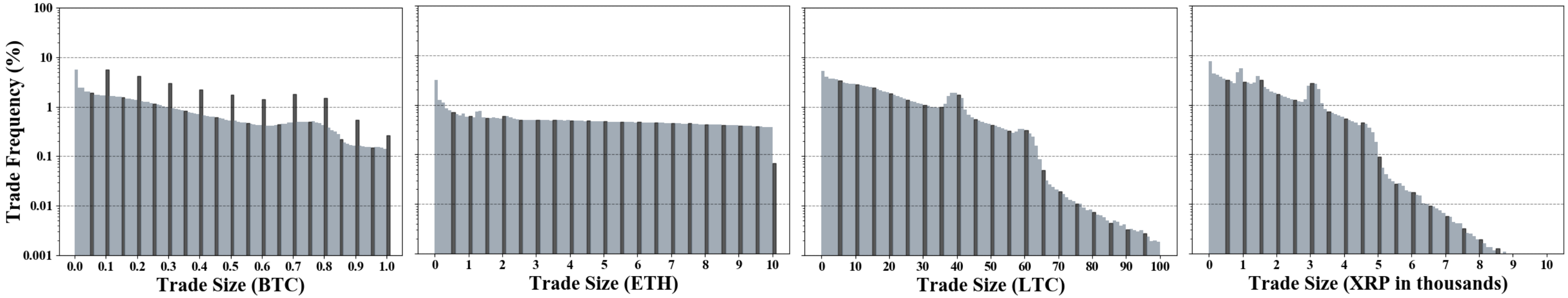

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (XRP in thousands)

U10

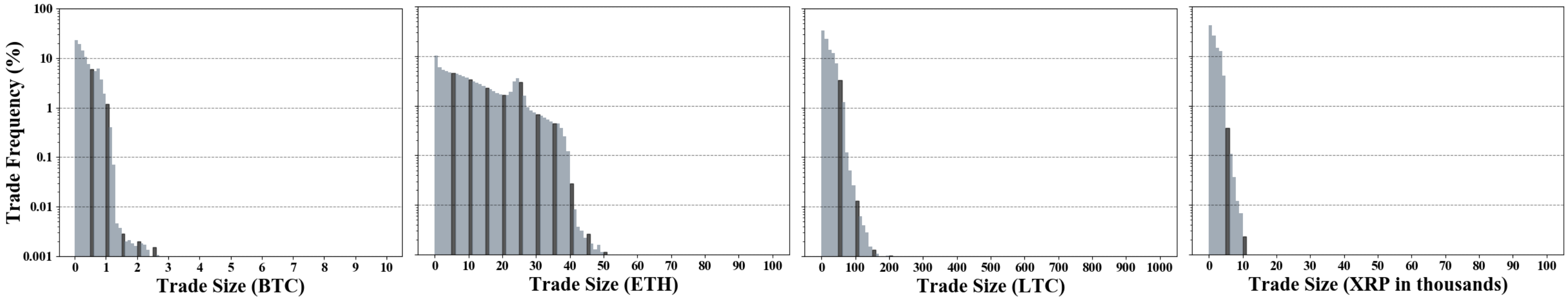


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)
Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0 1 2 3 4 5 6 7 8 9 10
Trade Size (BTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (ETH)
0 100 200 300 400 500 600 700 800 900 1000
Trade Size (LTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (XRP in thousands)

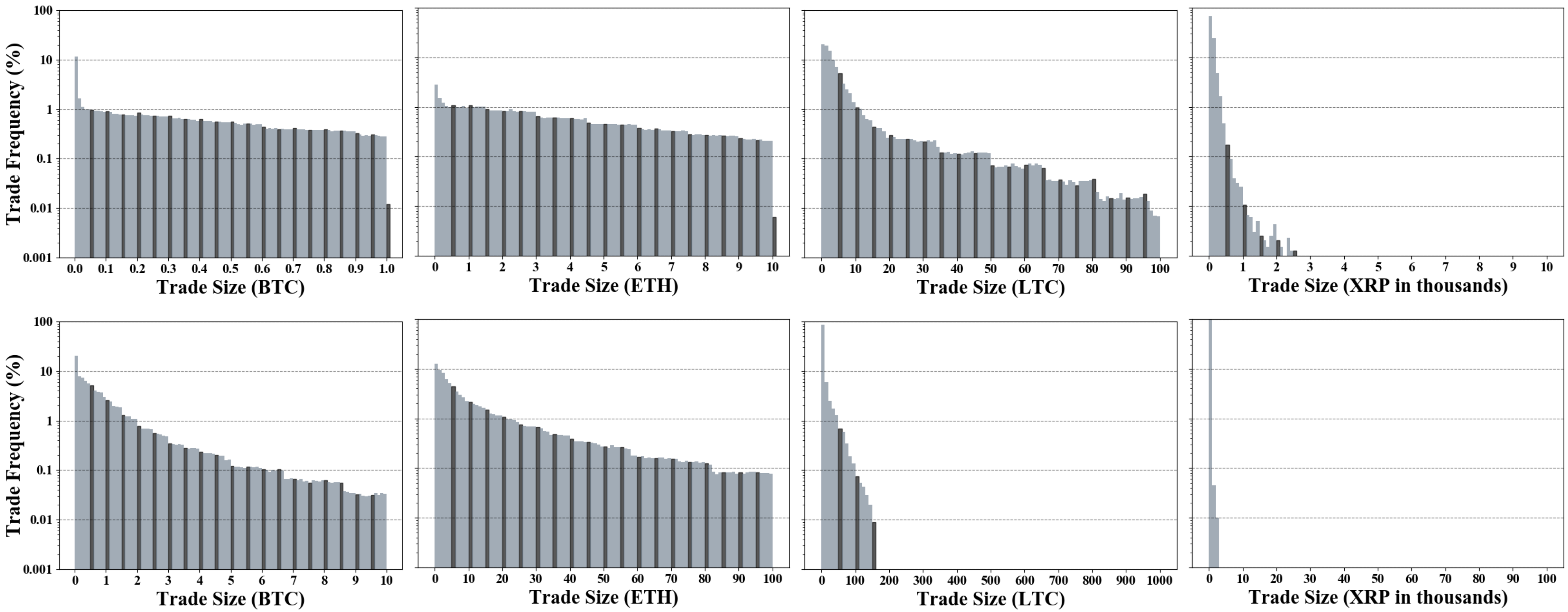
U11



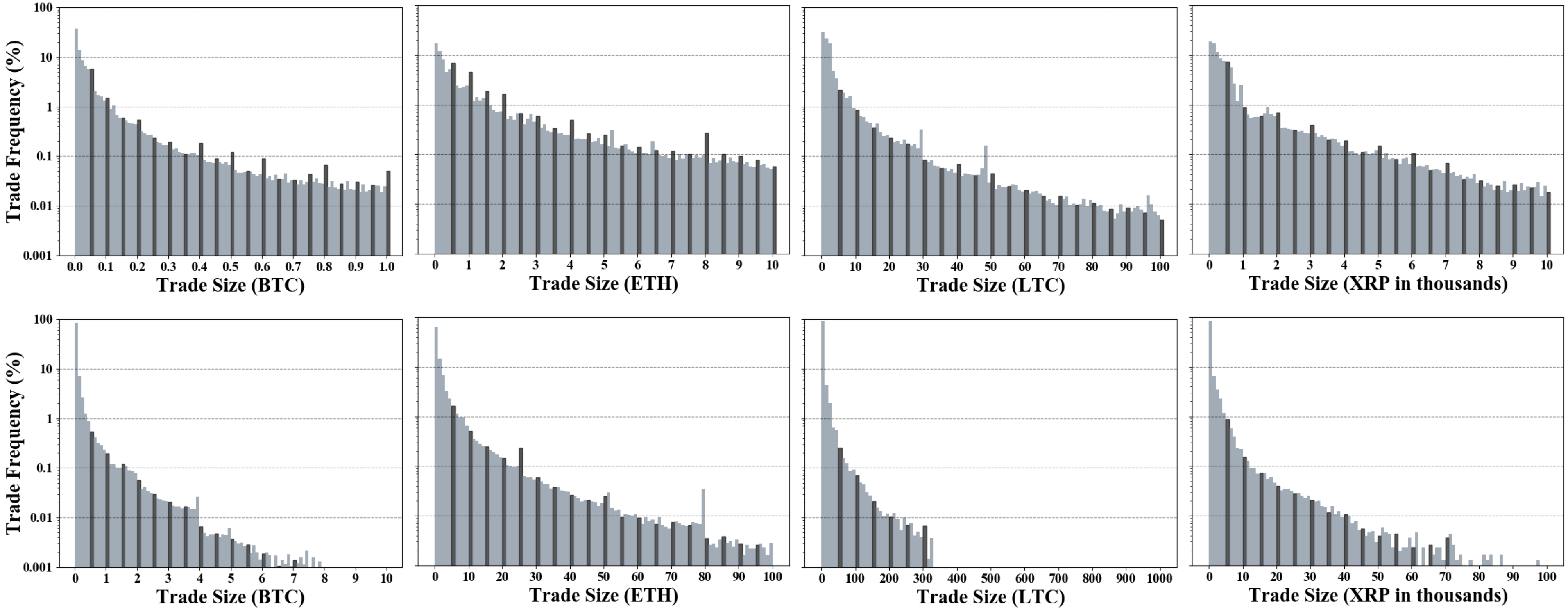
Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)


Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

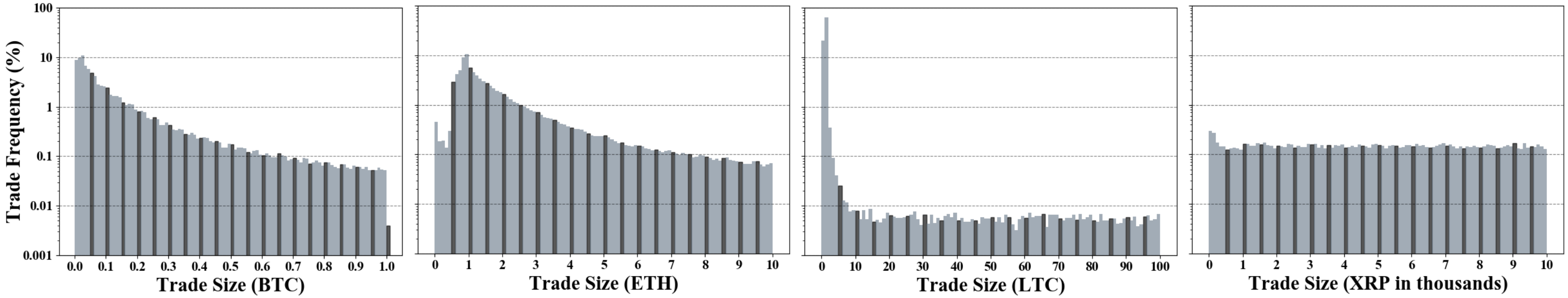
U12

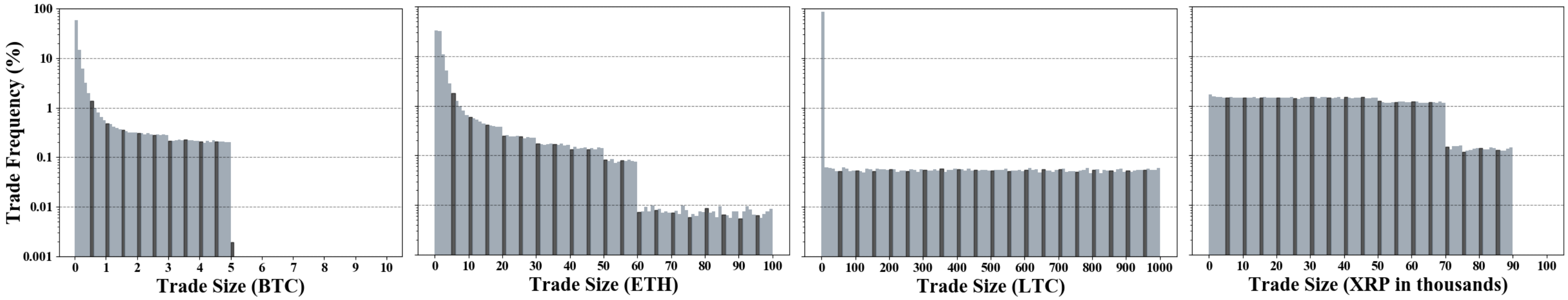

Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U13

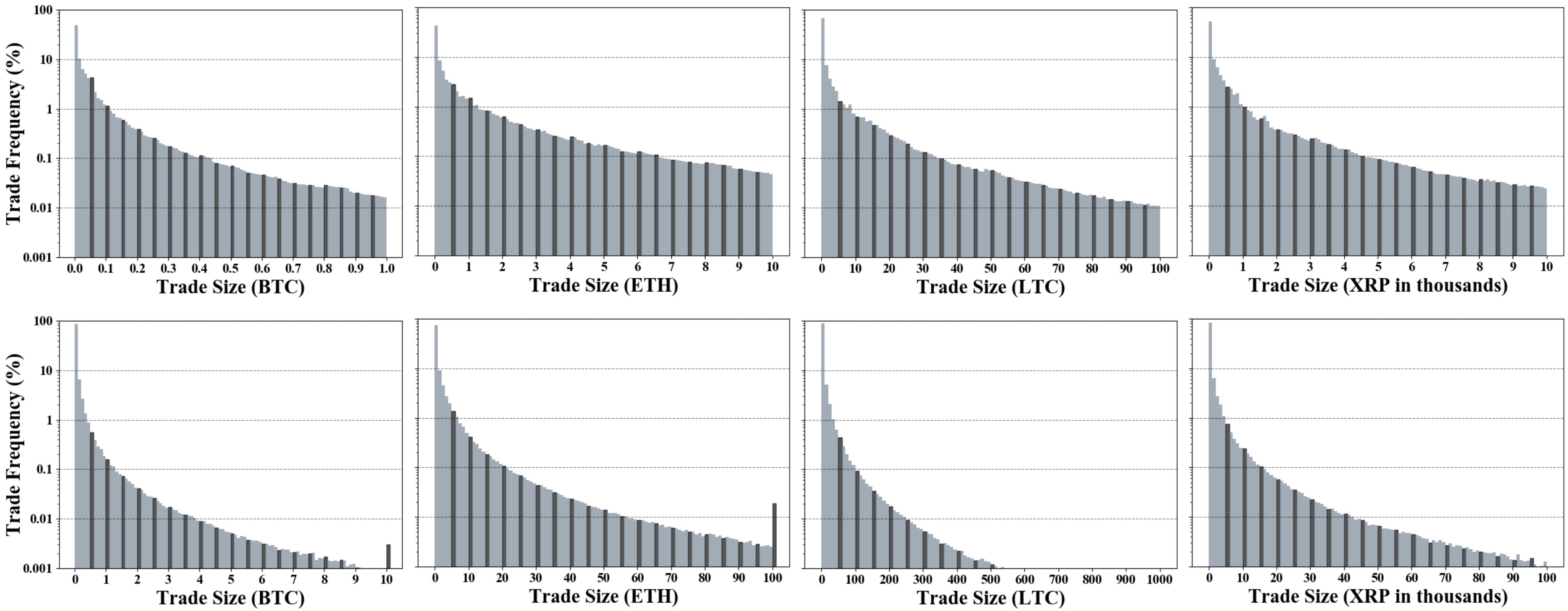

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

U14

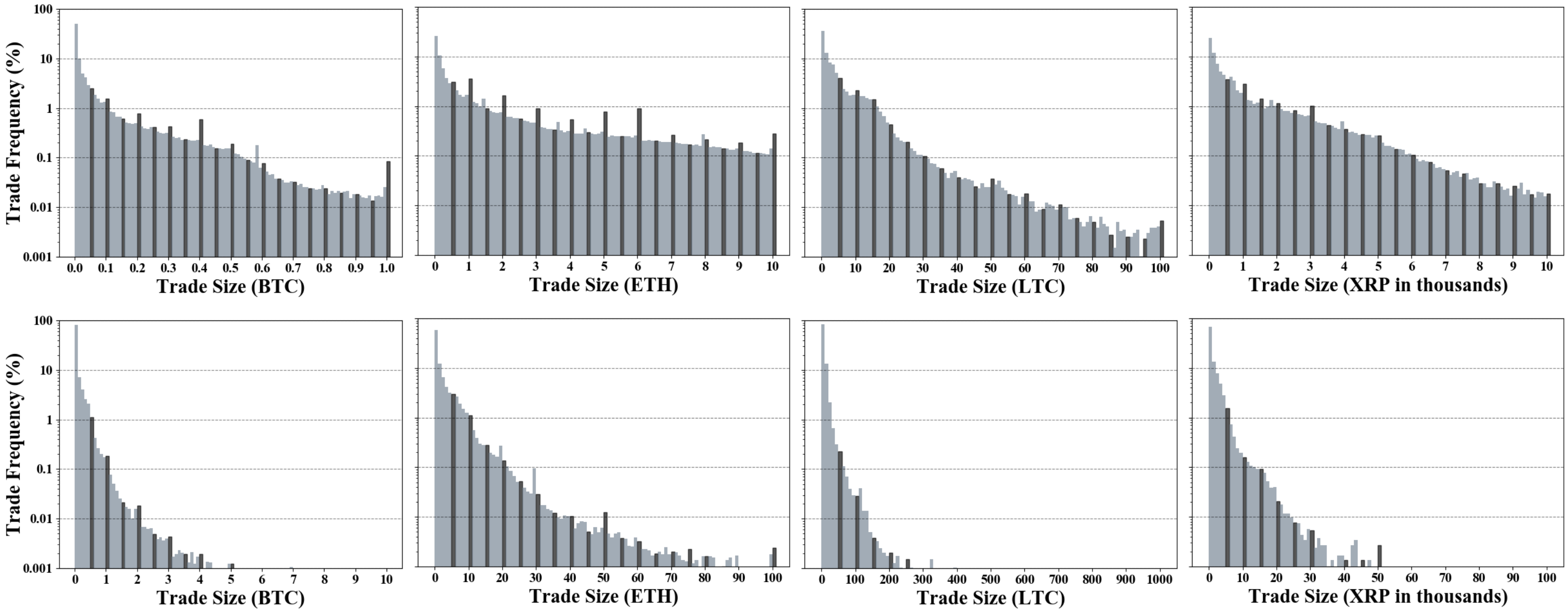

Trade Frequency (%)
100
10
1
0.1
0.01
0.001
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0 1 2 3 4 5 6 7 8 9 10
Trade Size (BTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (ETH)
0 100 200 300 400 500 600 700 800 900 1000
Trade Size (LTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (XRP in thousands)

U15


Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Trade Size (BTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (ETH)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (LTC)
0 1 2 3 4 5 6 7 8 9 10
Trade Size (XRP in thousands)
Trade Frequency (%)
100
10
1
0.1
0.01
0.001
0 1 2 3 4 5 6 7 8 9 10
Trade Size (BTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (ETH)
0 100 200 300 400 500 600 700 800 900 1000
Trade Size (LTC)
0 10 20 30 40 50 60 70 80 90 100
Trade Size (XRP in thousands)

U16


Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)
Trade Frequency (%)
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP in thousands)

Appendix C Regression analysis for trade-size clustering

Appendix C reports regression analysis for trade size-clustering effect. The following regression is estimated for each crypto exchange with Ordinary-Least-Square (OLS) method.

$$Logit\_Percentage_i = \alpha + \beta_{100} * M_{100,i} + \beta_{500} * M_{500,i} + \beta_{1000} * M_{1000,i} + \beta_{5000} * M_{5000,i} + \beta_{10000} * M_{10000,i} + \varepsilon_i$$

The dependent variable is $Logit\_Percentage_i$, is calculated via the logit function of the percentage of trades with size i in all trades of each crypto exchange. $M_{100,i}$ is a dummy variable which equals 1 if trade size is multiple of 100 units, but not the multiple of 500 units, 1000units, 5000units and 10,000 units for each cryptocurrency pair. $M_{500,i}$ is a dummy variable which equals 1 if trade size is multiple of 500 units, but not the multiple of 1000units, 5000units and 10,000 units for each cryptocurrency pair. $M_{1000,i}$ is a dummy variable which equals 1 if trade size is multiple of 1000 units, but not the multiple of 5000units and 10,000 units for each cryptocurrency pair. $M_{5000,i}$is a dummy variable which equals 1 if trade size is multiple of 5000 units but not the multiple of 10,000 units. $M_{10000,i}$ is a dummy variable which equals 1 if trade size is multiple of 10,000 units for each cryptocurrency pair. t-statistics are reported in the brackets. ***, ** and * represents the statistical significant level at 1%, 5% and 1%, respectively.

Panel R: Regulated Exchanges

Logit_Percentage	BTC/USD			ETH/USD			LTC/USD			XRP/USD	
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2
M_{100}	1.816***	1.750***	1.321***	1.078***	1.975***	0.700***	1.621***	2.060***	1.693***	1.202***	0.320**
	(25.55)	(24.31)	(17.18)	(18.16)	(29.49)	(10.84)	(15.33)	(13.39)	(11.53)	(17.85)	(2.54)
M_{500}	1.789***	2.001***	1.835***	1.301***	1.882***	1.164***	1.933***	2.487***	1.892***	1.175***	1.429***
	(8.70)	(7.96)	(6.37)	(7.30)	(9.36)	(5.56)	(5.07)	(5.21)	(3.72)	(6.65)	(3.18)
M_{1000}	1.896***	1.585***	1.671***	1.735***	2.321***	1.891***	1.681***	2.349***	2.073***	1.351***	1.399***
	(12.00)	(11.03)	(9.14)	(12.77)	(15.31)	(10.83)	(6.63)	(7.54)	(5.48)	(8.25)	(3.27)
M_{5000}	2.549***	1.761***	2.087***	2.176***	2.551***	2.914***	1.894***	2.310***	4.065***	1.808***	1.832*
	(4.79)	(3.74)	(3.25)	(4.63)	(5.28)	(4.26)	(2.75)	(3.02)	(4.25)	(4.06)	(1.65)
M_{1000}	2.523***	2.400***	2.357***	2.313***	3.259***	2.670***	1.851**	2.591***	1.746**	1.130***	0.554
	(6.44)	(5.14)	(4.67)	(5.97)	(8.01)	(6.27)	(2.42)	(3.64)	(2.05)	(7.02)	(0.96)
Constant	-14.404***	-15.324***	-13.039***	-12.819***	-14.284***	-11.967***	-12.037***	-14.043***	-11.072***	-13.077***	-12.671***
	(-32819.11)	(-50870.44)	(-20439.31)	(-13446.87)	(-41185.85)	(-11209.64)	(-8640.73)	(-33087.58)	(-4694.29)	(-17555.21)	(-1241.56)
Observations	1081863	2991928	303572	214912	1016338	111327	98604	789721	40324	259247	26380
Adjusted R2	0.017	0.004	0.036	0.040	0.035	0.062	0.050	0.014	0.055	0.032	0.003

Panel UT: Unregulated Tier-1 Exchanges

Logit_Percentage	BTC/USD									
	UT1	UT2	UT3	UT4	UT5	UT6	UT7	UT8	UT9	UT10
M_{100}	0.690***	1.725***	1.991***	1.157***	1.759***	2.072***	1.723***	2.402***	1.823***	0.130
	(6.81)	(11.55)	(19.40)	(10.27)	(17.72)	(6.14)	(11.75)	(18.33)	(17.21)	(1.44)
M_{500}	1.008***	2.121***	2.496***	1.392***	2.044***	2.252**	1.970***	2.493***	1.874***	0.232
	(3.14)	(4.35)	(7.16)	(4.17)	(6.67)	(2.47)	(4.40)	(6.17)	(6.08)	(0.88)
M_{1000}	1.425***	2.699***	2.009***	2.053***	1.760***	3.468***	2.268***	2.157***	2.210***	0.175
	(9.76)	(3.86)	(9.72)	(4.61)	(12.99)	(3.05)	(5.41)	(11.47)	(5.51)	(0.71)
M_{5000}	1.306***	1.786	1.998***	2.423***	2.145***	4.551***	2.648**	1.924***	2.521**	-0.093
	(2.90)	(1.48)	(3.73)	(2.62)	(5.21)	(10420.69)	(2.33)	(3.69)	(2.33)	(-0.14)
M_{1000}	0.408	3.147***	2.246***	3.291***	2.129***	1.184***	3.458***	2.204***	2.812***	-0.474
	(1.54)	(3.56)	(5.97)	(3.30)	(5.58)	(2710.76)	(4.49)	(4.16)	(4.59)	(-0.99)
Constant	-15.879***	-12.601***	-15.080***	-14.352***	-16.561***	-14.699***	-12.548***	-16.098***	-13.184***	-14.751***
	(-10918.52)	(-16883.65)	(-36056.36)	(-7850.53)	(-21698.85)	(-33655.68)	(-8334.49)	(-50580.53)	(-17332.25)	(-1963.55)
Observations	1133210	218660	1328575	293015	2023008	1377608	86476	5237865	368634	60972
Adjusted R2	0.000	0.025	0.011	0.003	0.001	0.001	0.053	0.001	0.018	-0.000

Logit_Percentage	ETH/USD									
	UT1	UT2	UT3	UT4	UT5	UT6	UT7	UT8	UT9	UT10
M_{100}	1.200***	1.359***	1.432***	0.616***	1.323***	2.658***	0.961***	1.615***	1.030***	0.006
	(12.19)	(8.21)	(17.70)	(7.43)	(16.05)	(7.95)	(8.40)	(15.66)	(12.55)	(0.07)
M_{500}	1.598***	1.806***	2.120***	1.337***	1.390***	2.943***	1.583***	1.791***	1.287***	-0.029
	(5.05)	(3.48)	(7.59)	(3.79)	(5.68)	(3.68)	(3.33)	(5.14)	(4.47)	(-0.12)
M_{1000}	1.165***	2.100***	2.042***	1.636***	1.068***	3.883***	1.766***	1.655***	1.696***	0.456
	(11.25)	(5.15)	(11.56)	(6.35)	(9.85)	(3.92)	(5.72)	(10.10)	(8.37)	(1.40)
M_{5000}	1.572***	2.753**	2.926***	2.566***	1.296***	4.568***	1.291	1.817***	2.231***	0.254
	(4.78)	(2.23)	(4.56)	(2.61)	(3.72)	(14670.10)	(1.54)	(3.48)	(3.59)	(0.29)
M_{1000}	1.313***	1.855***	2.662***	2.229***	1.772***	5.474***	2.939***	1.941***	2.107***	1.526**
	(4.20)	(3.35)	(6.22)	(3.29)	(4.95)	(3.24)	(4.95)	(4.33)	(4.39)	(1.98)
Constant	-15.209***	-11.588***	-14.190***	-14.084***	-15.672***	-14.717***	-11.198***	-15.862***	-12.255***	-13.660***
	(-14440.78)	(-12133.85)	(-32615.23)	(-9884.54)	(-11234.33)	(-47268.36)	(-5134.03)	(-28788.94)	(-14299.09)	(-1861.97)
Observations	1184422	83685	787968	359258	825286	1454571	42352	2358413	156313	64035
Adjusted R2	0.001	0.062	0.021	0.002	0.002	0.004	0.051	0.001	0.035	-0.000

Logit_Percentage	LTC/USD								
	UT1	UT2	UT3	UT4	UT5	UT6	UT8	UT9	UT10
M_{100}	2.345***	1.610***	1.972***	0.895***	1.502***	3.504***	2.485***	1.591***	0.049
	(11.48)	(8.06)	(10.48)	(6.08)	(7.38)	(5.09)	(10.18)	(8.13)	(0.32)
M_{500}	2.704***	1.841***	2.692***	1.355**	1.546***	2.696*	2.403***	2.876***	0.016
	(4.26)	(3.13)	(5.15)	(2.29)	(2.64)	(1.85)	(2.93)	(4.24)	(0.04)
M_{1000}	2.048***	2.244***	2.569***	1.858***	1.402***	3.388*	1.586***	1.647***	0.389
	(7.55)	(2.69)	(6.64)	(2.74)	(5.25)	(1.80)	(5.00)	(3.83)	(0.82)
M_{5000}	2.129***	4.562***	2.075**	2.062	1.642***	2.293***	1.266**	2.427**	0.905***
	(2.95)	(2757.10)	(2.50)	(1.57)	(2.66)	(7087.76)	(1.99)	(2.44)	(247.47)
M_{1000}	1.939***	1.986***	2.115***	1.077	1.953***		1.443**	1.766**	
	(3.36)	(4.05)	(3.13)	(1.64)	(3.33)		(2.39)	(2.17)	
Constant	-14.611***	-10.797***	-13.381***	-13.032***	-14.488***	-14.339***	-15.241***	-11.180***	-14.068***
	(-17060.06)	(-6525.53)	(-21679.02)	(-5904.67)	(-9372.61)	(-44322.12)	(-29947.04)	(-8059.40)	(-3847.76)
Observations	880422	39020	338063	145449	486726	1170269	1997860	57379	146656
Adjusted R2	0.003	0.053	0.027	0.003	0.001	0.002	0.001	0.039	-0.000

Logit_Percentage	XRP/USD									
	UT1	UT2	UT3	UT4	UT5	UT6	UT7	UT8	UT9	UT10
M_{100}	0.739***	1.400***	1.532***	0.722***	1.658***	3.395***	1.145***	2.060***	1.309***	0.080
	(4.56)	(8.17)	(12.43)	(4.69)	(8.77)	(6.83)	(7.53)	(8.28)	(7.23)	(0.97)
M_{500}	1.620***	1.689***	2.237***	1.322**	2.249***	3.304***	1.520***	2.613***	1.326***	0.073
	(3.08)	(4.02)	(5.79)	(2.26)	(3.68)	(2.58)	(2.60)	(3.53)	(2.76)	(0.34)
M_{1000}	1.241**	1.450***	2.257***	2.366***	1.795***	3.627***	1.866***	2.246***	1.476***	-0.012
	(2.54)	(4.02)	(5.85)	(3.77)	(3.97)	(3.75)	(5.11)	(4.69)	(4.59)	(-0.05)
M_{5000}	2.553**	2.738**	3.138***	4.276***	2.365**	1.721***	2.602***	2.453**	2.054*	0.028
	(2.50)	(2.37)	(2.97)	(3.25)	(2.26)	(6982.17)	(3.53)	(2.01)	(1.71)	(0.05)
M_{1000}	1.706**	1.083**	2.160***	2.913**	1.893***	-0.070***	1.950***	1.764*	0.683***	-0.119
	(2.20)	(2.45)	(15.76)	(2.14)	(2.68)	(-285.37)	(3.61)	(1.71)	(2.74)	(-0.29)
Constant	-14.051***	-10.754***	-13.786***	-12.894***	-14.713***	-14.537***	-10.343***	-14.754***	-11.605***	-13.416***
	(-3504.97)	(-7767.92)	(-26701.42)	(-6002.19)	(-14382.69)	(-58964.68)	(-3097.99)	(-20595.30)	(-11564.37)	(-1902.72)

Observations	161287	39993	375744	137671	853194	1525733	16867	1142736	76975	57810
Adjusted R2	0.001	0.072	0.043	0.003	0.001	0.004	0.091	0.001	0.036	-0.000

Panel U: Unregulated Tier-2 Exchanges

Logit_Percentage	BTC/USD															
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15	U16
M_{100}	0.118	1.652***	-0.052	0.046	0.662***	-0.017	2.367***	0.841***	1.524***	0.009	-0.016	0.428***	1.398***	1.031***	0.184**	2.702***
	(1.41)	(18.24)	(-0.43)	(0.41)	(6.69)	(-0.46)	(10.90)	(13.00)	(2.61)	(0.09)	(-0.06)	(7.05)	(12.83)	(30.77)	(2.15)	(15.39)
M_{500}	-0.192	1.675***	-0.159	0.306	0.729***	-0.005	2.446***	0.956***	2.121	0.021	-0.168	0.542***	2.121***	1.053***	0.228	2.973***
	(-1.44)	(6.98)	(-0.46)	(0.96)	(2.71)	(-0.05)	(3.86)	(5.28)	(1.30)	(0.10)	(-0.24)	(3.08)	(6.20)	(12.23)	(0.90)	(5.64)
M_{1000}	0.174	1.978***	0.094	0.349	0.935***	-0.021	2.274***	1.219***	2.571***	-0.014	1.180	1.036***	2.522***	1.003***	0.124	2.709***
	(0.63)	(6.77)	(0.24)	(0.92)	(2.86)	(-0.18)	(3.79)	(5.75)	(3.36)	(-0.07)	(1.55)	(5.84)	(7.97)	(10.54)	(0.53)	(4.67)
M_{5000}	0.676***	2.323***	-0.443	0.605	1.833	0.506	1.990	2.009***	0.296***	-0.378***	0.410	1.228**	3.449***	0.915***	0.088	4.031***
	(1001.83)	(4.82)	(-0.73)	(0.72)	(1.54)	(1.39)	(1.20)	(4.08)	(90.45)	(-35.06)	(0.21)	(2.52)	(3.44)	(7.54)	(0.14)	(2.86)
M_{1000}	.	2.721***	-0.352	-1.207	-0.967*	-0.387	7.165***	2.760***	1.213***		0.481	0.884**	2.553***	0.957***	0.736	3.063***
	.	(4.07)	(-0.25)	(-1.57)	(-1.65)	(-1.42)	(4.66)	(6.71)	(370.07)		(0.23)	(2.01)	(3.43)	(7.19)	(1.19)	(4.20)
Constant	-13.482***	-14.081***	-12.073***	-12.379***	-12.800***	-12.784***	-14.490***	-16.869***	-11.807***	-9.074***	-11.326***	-13.113***	-14.082***	-11.515***	-15.004***	-14.106***
	(-19985.05)	(-27545.69)	(-782.47)	(-1289.09)	(-1451.11)	(-3784.47)	(-38643.56)	(-14371.34)	(-3602.97)	(-842.03)	(-650.48)	(-8971.58)	(-11030.44)	(-973.48)	(-2006.09)	(-11720.98)
Observations	612213	905078	18300	44521	39658	142678	1586498	406726	68548	5496	14416	309459	444827	19947	64026	274884
Adjusted R2	0.000	0.004	-0.000	-0.000	0.002	-0.000	0.004	0.004	0.001	-0.001	0.000	0.001	0.004	0.009	0.000	0.022

Logit_Percentage	ETH/USD															
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15	U16
M_{100}	0.487***	0.873***	1.620***	2.252***	1.358***	0.108***	1.717***	1.319***	1.590***	0.018	2.039***	1.880***	1.398***	-0.218***	0.694***	2.283***
	(3.21)	(24.41)	(7.61)	(19.48)	(12.13)	(4.75)	(15.66)	(13.17)	(3.32)	(0.13)	(14.63)	(28.21)	(14.64)	(-4.86)	(8.85)	(18.99)
M_{500}	-0.012***	0.899***	1.077**	2.408***	0.952***	0.123	2.211***	1.325***	1.471	0.272	1.975***	1.671***	1.679***	-0.138	0.648***	2.018***
	(-128.34)	(8.54)	(2.12)	(6.94)	(3.20)	(1.58)	(6.12)	(4.34)	(1.52)	(0.85)	(4.95)	(8.17)	(5.92)	(-0.89)	(2.84)	(5.99)
M_{1000}	1.028***	1.164***	2.003***	1.970***	1.948***	0.248**	6.021***	1.671***	2.514**	-0.150	1.679***	1.655***	1.851***	-0.040	0.911***	2.364***
	(4.19)	(6.71)	(2.70)	(4.52)	(5.34)	(2.10)	(18.22)	(5.28)	(2.05)	(-1.15)	(4.01)	(6.72)	(6.79)	(-0.24)	(3.37)	(6.39)
M_{5000}	-0.012***	1.618***	1.679	3.166***	0.217	0.549	6.161***	1.872**	0.725***	.	2.409*	2.097***	2.010***	-0.311	0.627	2.629***
	(-128.34)	(3.17)	(1.14)	(3.11)	(0.21)	(1.51)	(6.28)	(2.30)	(211.16)	.	(1.84)	(3.84)	(3.17)	(-0.98)	(0.94)	(2.72)
M_{1000}	.	2.032***	0.612	0.432	0.499	0.133	3.425***	1.940***	-0.374***	.	1.952	1.447**	2.190***	-0.538**	1.168	1.984***
	.	(3.86)	(0.94)	(0.53)	(0.72)	(0.48)	(3.82)	(4.90)	(-108.81)	.	(1.50)	(2.11)	(3.92)	(-2.23)	(1.48)	(3.88)

Constant	-13.884***	-13.598***	-13.113***	-13.524***	-13.547***	-13.704***	-14.657***	-16.803***	-11.778***	-10.804***	-13.782***	-13.682***	-13.504***	-12.506***	-15.217***	-13.736***
	(-151428.17)	(-16607.11)	(-3680.38)	(-5581.38)	(-5001.16)	(-23408.55)	(-55985.57)	(-13509.20)	(-3430.37)	(-2453.64)	(-7955.51)	(-36188.51)	(-11300.09)	(-3799.39)	(-5969.91)	(-11068.89)
Observations	1063669	590814	114107	289400	225927	750844	1879891	677414	64043	32855	362960	806140	285741	116774	373622	313714
Adjusted R2	0.001	0.003	0.004	0.004	0.001	0.000	0.017	0.002	0.002	-0.000	0.008	0.015	0.011	0.000	0.000	0.014

Logit_Percentage	LTC/USD												
	U2	U3	U4	U5	U7	U8	U10	U11	U12	U13	U14	U15	U16
M_{100}	0.509***	-0.591	0.349	1.548***	5.211***	1.501***	0.070	2.050***	-0.002	2.631***	-0.185**	0.623***	1.868***
	(6.24)	(-1.25)	(1.27)	(4.71)	(14.02)	(3.99)	(0.62)	(8.00)	(-0.01)	(8.36)	(-2.44)	(3.03)	(7.27)
M_{500}	0.513**	0.354***	1.700*	3.860***	4.243***	1.469	0.095***	1.828***	-0.121	2.428***	-0.347***	0.331	2.528***
	(2.24)	(67.18)	(1.79)	(3.32)	(3.60)	(1.40)	(30.74)	(2.75)	(-0.20)	(2.76)	(-98.21)	(0.78)	(2.98)
M_{1000}	0.777***	-2.092***	0.560	2.254**	5.293***	1.863***	.	1.490**	-0.474	1.712***	-0.270***	0.563	3.218***
	(2.91)	(-5.39)	(0.91)	(2.02)	(4.06)	(3.05)	.	(2.13)	(-1.49)	(3.06)	(-3.72)	(1.04)	(4.27)
M_{5000}	0.299	.	-0.563***	-0.120***	6.310***	0.836	.	2.939	-1.071***	1.919	-0.347***	-0.287***	3.695***
	(0.77)	.	(-283.03)	(-66.37)	(6625.23)	(0.80)	.	(1.03)	(-167.99)	(1.27)	(-98.21)	(-166.00)	(2329.59)
M_{1000}	-0.250***	.	-1.661***	.	-1.264***	1.793***	.	0.515	-0.522	1.572**	.	-0.980***	1.383*
	(-479.40)	.	(-835.60)	.	(-1326.88)	(2.64)	.	(0.78)	(-1.34)	(2.02)	.	(-566.54)	(1.70)
Constant	-13.511***	-11.529***	-13.891***	-13.989***	-13.529***	-16.135***	-11.317***	-14.186***	-12.092***	-12.522***	-13.026***	-14.997***	-12.642***
	(-25873.02)	(-2187.40)	(-6986.98)	(-7722.10)	(-14204.16)	(-6797.88)	(-3677.56)	(-11368.51)	(-1896.05)	(-6987.14)	(-3683.50)	(-8666.15)	(-7970.32)
Observations	659641	43382	483967	317024	541896	365449	54554	545009	47228	112788	102704	522475	158071
Adjusted R2	0.000	0.000	0.000	0.001	0.008	0.000	-0.000	0.002	-0.000	0.015	-0.000	0.000	0.007

Logit_Percentage	XRP/USD													
	U2	U3	U4	U5	U7	U8	U9	U10	U11	U12	U13	U14	U15	U16
M_{100}	0.484***	2.411***	0.269	2.812***	0.465***	1.181***	1.321**	0.195	3.069***	-0.704	1.845***	0.797***	1.090***	1.863***
	(3.02)	(3.00)	(0.94)	(6.34)	(4.76)	(4.12)	(2.20)	(0.80)	(11.16)	(-1.30)	(10.27)	(2.58)	(5.70)	(5.97)
M_{500}	1.299**	0.823***	0.574	2.708	0.726*	1.153*	2.336	.	3.536***	.	2.011***	1.791***	1.257***	2.284***
	(2.28)	(1156.26)	(0.91)	(1.54)	(1.92)	(1.65)	(1.16)	.	(4.67)	.	(5.55)	(3.65)	(3.36)	(2.83)
M_{1000}	1.593***	1.333***	-0.000	2.411***	0.990**	1.093***	0.176***	.	3.723***	-1.625***	2.101***	2.078***	2.463***	2.488***
	(2.93)	(1874.25)	(-0.00)	(7.93)	(2.45)	(2.69)	(37.38)	.	(4.89)	(-231.27)	(6.39)	(24424.40)	(4.52)	(3.28)
M_{5000}	1.081	.	-0.367***	.	2.041***	1.593**	.	.	1.634***	-0.932***	2.085**	.	0.605	1.881**
	(1.08)	.	(-185.13)	.	(2.70)	(2.19)	.	.	(4900.47)	(-132.64)	(2.10)	.	(0.74)	(2.00)
M_{1000}	0.816	.	-1.465***	.	0.234	1.518***	-0.206	.	.	-1.625***	1.199**	.	-0.119	1.015***
	(1.00)	.	(-739.97)	.	(0.44)	(3.83)	(-1.07)	.	.	(-231.27)	(2.35)	.	(-0.31)	(3.21)

Constant	-13.155***	-13.971***	-13.683***	-14.153***	-14.323***	-15.934***	-11.505***	-14.210***	-15.000***	-11.253***	-12.431***	-11.960***	-15.314***	-12.533***
	(-17194.82)	(-19636.73)	(-6910.18)	(-15273.21)	(-35359.12)	(-7994.80)	(-2439.92)	(-52628.48)	(-44989.61)	(-1601.34)	(-5198.38)	(-140549.19)	(-23502.14)	(-19122.97)
Observations	441032	723474	426805	678851	1453660	387182	44950	1395135	2209056	35362	78777	156112	1879104	212728
Adjusted R2	0.000	0.000	-0.000	0.001	0.000	0.000	0.001	-0.000	0.001	0.000	0.021	0.082	0.000	0.017

Appendix D Chi-square tests for trade-size distribution

Appendix D presents the Pearson's Chi-square statistics and test the null hypothesis that trade-size distributions of unregulated exchanges are not significantly different from those of regulated exchanges. Distributions of four trading pairs are reported, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank. χ^2 statistics and *p-value* are reported in the Table. ***, ** and * represents the statistical significant level at 1%, 5% and 1%, respectively.

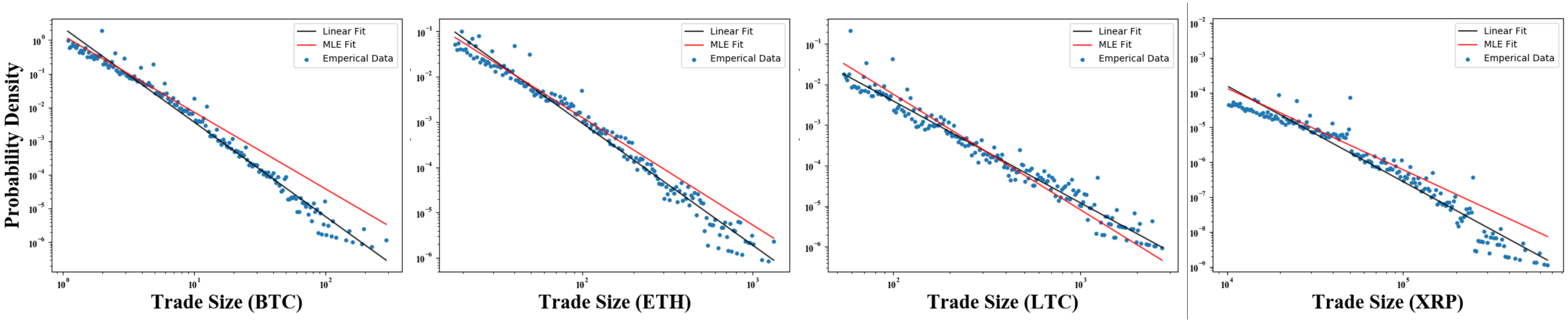
Exchange Code	BTC/USD		ETH/USD		LTC/USD		XRP/USD	
	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value
Panel A Regulated exchanges								
R1	3.101	0.541	6.353	0.174	7.554	0.109	9.238	0.161
R2	1.954	0.744	3.631	0.458	3.073	0.546	9.238	0.161
R3	1.227	0.874	0.898	0.925	1.682	0.794		
Panel B Unregulated Tier-1 exchanges								
UT1	11.923**	0.018	9.962**	0.041	14.354***	0.006	15.686**	0.016
UT2	5.412	0.248	9.709**	0.046	15.991***	0.003	9.942	0.127
UT3	1.791	0.774	2.367	0.669	1.942	0.746	16.210**	0.013
UT4	20.143***	0.000	38.374***	0.000	107.019***	0.000	121.556***	0.000
UT5	10.408**	0.034	10.69**	0.030	19.729***	0.001	37.646***	0.000
UT6	19.892***	0.001	34.822***	0.000	23.014***	0.000	160.611***	0.000
UT7	21.563***	0.000	1.636	0.802		0.000	7.208	0.302
UT8	8.15*	0.086	47.99***	0.000	44.178***	0.000	194.237***	0.000
UT9	7.194	0.126	15.744***	0.003	29.36***	0.010	141.321***	0.000
UT10	6.990	0.136	6.159	0.188	13.252***		20.665***	0.002
Panel C Unregulated Tier-2 exchanges								
U1	190.744***	0.000	214.171***	0.000				
U2	263.011***	0.000	300.510***	0.000	626.172***	0.000	40.188***	0.000
U3	76.301***	0.000	22.108***	0.000	58.221***	0.000	83.861***	0.000
U4	178.900***	0.000	48.256***	0.000	28.939***	0.000	19.924***	0.003
U5	297.113***	0.000	2.177	0.703	22.922***	0.000	35.933***	0.000
U6	605.266***	0.000	1062.603***	0.000				
U7	277.236***	0.000	345.818***	0.000	174.085***	0.000	335.128***	0.000
U8	78.193***	0.000	98.470***	0.000	129.744***	0.000	1205.640***	0.000
U9	81.483***	0.000	112.158***	0.000			83.003***	0.000
U10	39.108***	0.000	22.141***	0.000	42.896***	0.000	71.461***	0.000
U11	152.557***	0.000	116.467***	0.000	72.211***	0.000	46.129***	0.000
U12	164.513***	0.000	87.650***	0.000	41.086***	0.000	158.458***	0.000
U13	15.24***	0.004	18.157***	0.001	22.870***	0.000	26.642***	0.000
U14	45.314***	0.000	10.334**	0.035	2407.468***	0.000	1124.355***	0.000
U15	9.797**	0.044	32.081***	0.000	46.379***	0.000	61.833***	0.000
U16	9.052*	0.060	12.251**	0.016	9.181*	0.057	9.282	0.158

Appendix E Power Law

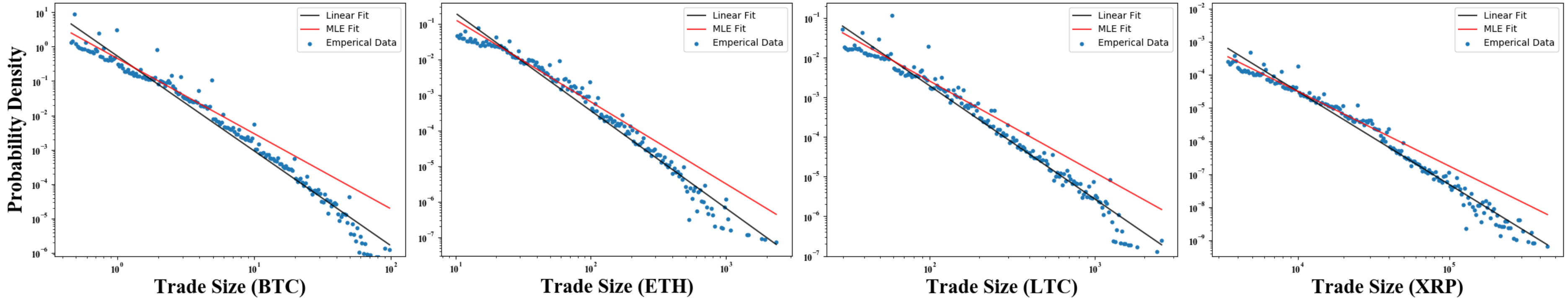
Appendix E display tail distribution of trade size and fitted power-law line on log-log scale. For each crypto exchange, four trading pairs are presented, including BTC/USD, ETH/USD, LTC/USD and XRP/USD. Y-axis represents the frequency of trades on a log scale while X-axis represents the logarithm of trade size. Fitted power-law lines are estimated by Ordinary Least Square (OLS) and Maximum Likelihood Estimation (MLE), plotted as black solid lines and red lines, respectively. Blue dots represent empirical data points for trade size. Panel R, Panel UT, and Panel U show distribution of trade-size in regulated exchanges, Tier-1 unregulated and Tier-2 unregulated exchanges, respectively. Regulated exchanges are those who are certified and regulated by New York State Department of Financial Services. Unregulated exchanges are categorized into unregulated Tier-1 and unregulated Tier-2 exchanges based on reputation and website traffic rank.

Panel R: Regulated exchange

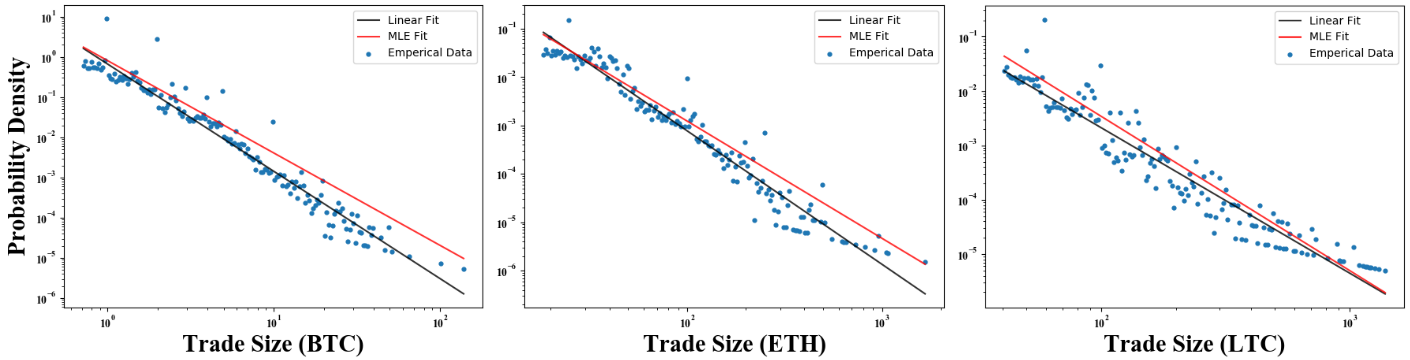
R1



R2

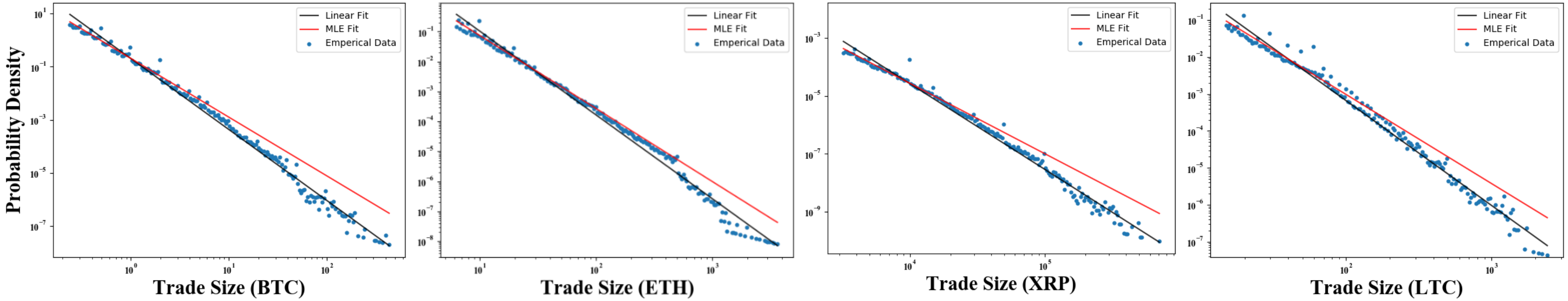


R3

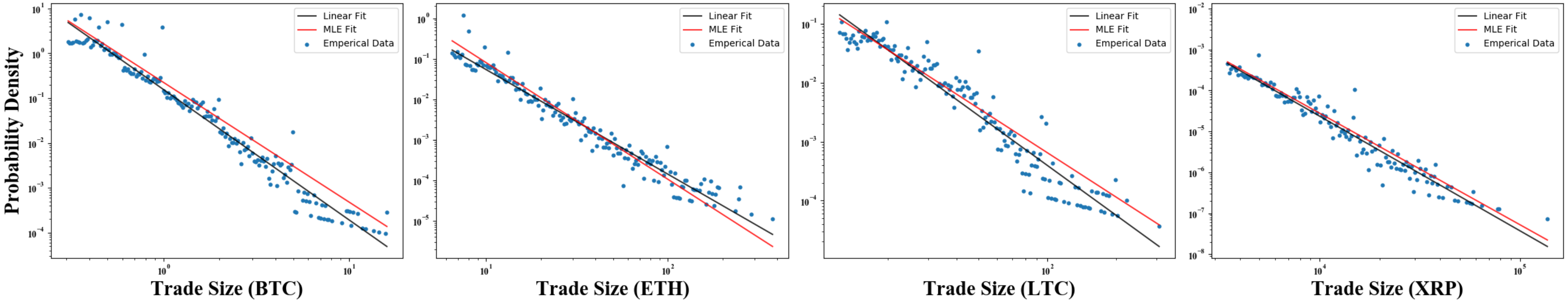


Panel UT: Unregulated Tier-1 exchanges

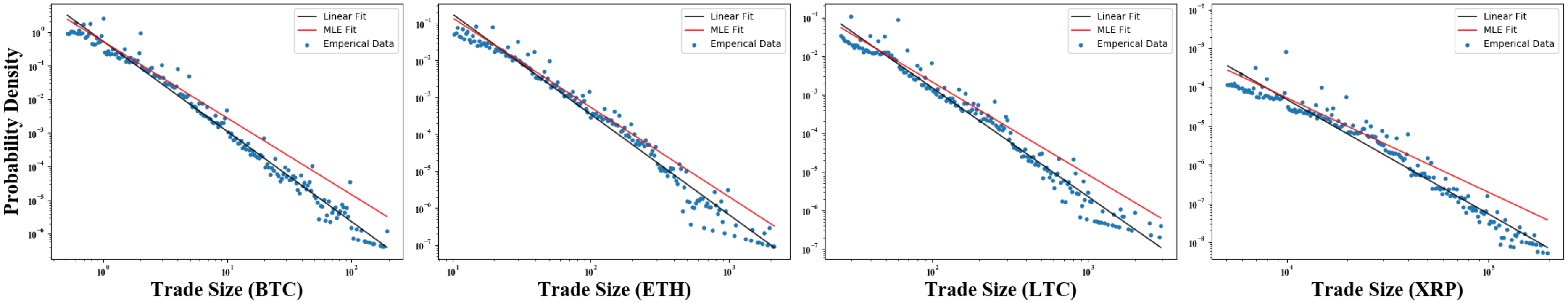
UT1


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)

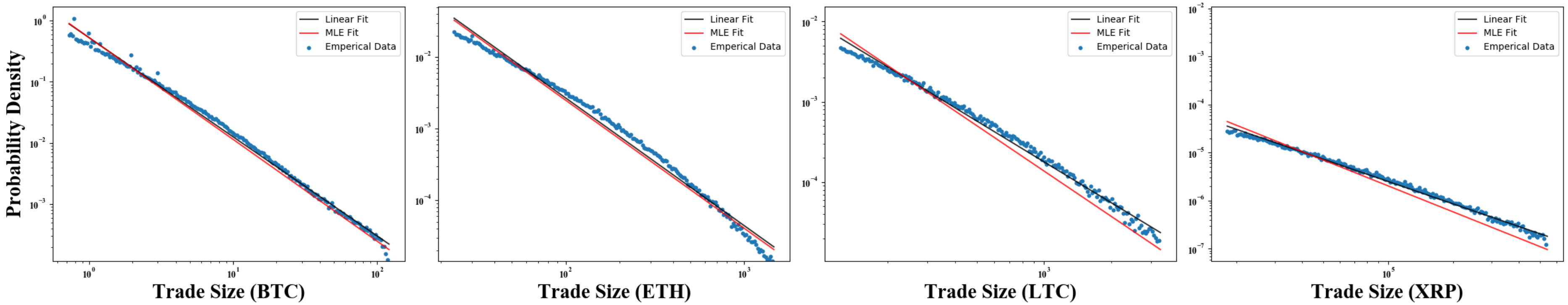
UT2


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

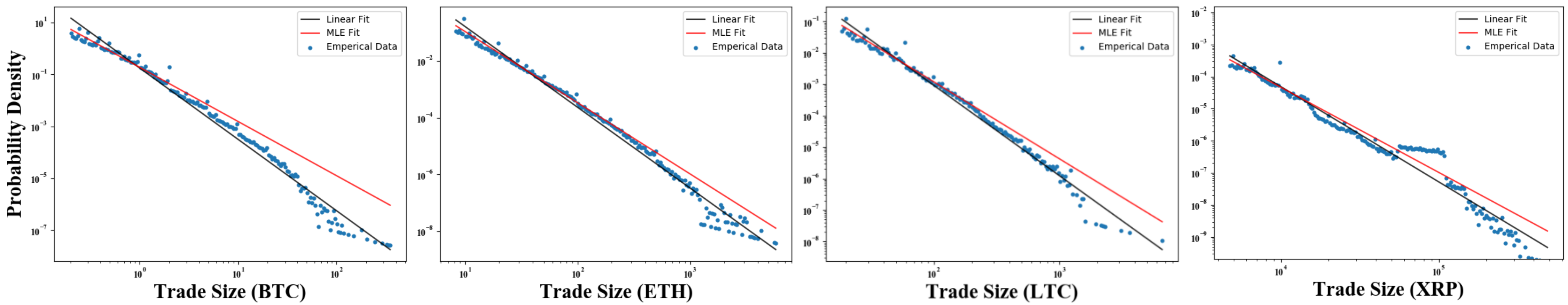
UT3


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

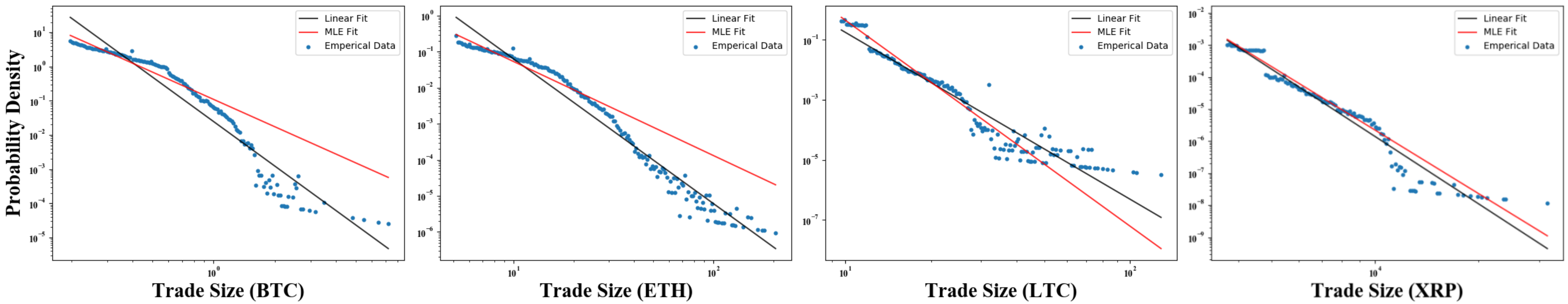
UT4


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

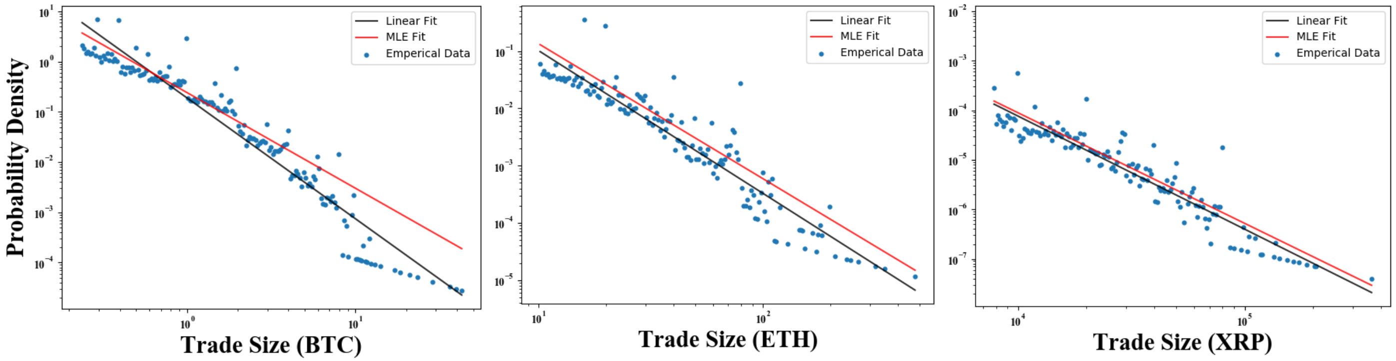
UT5


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

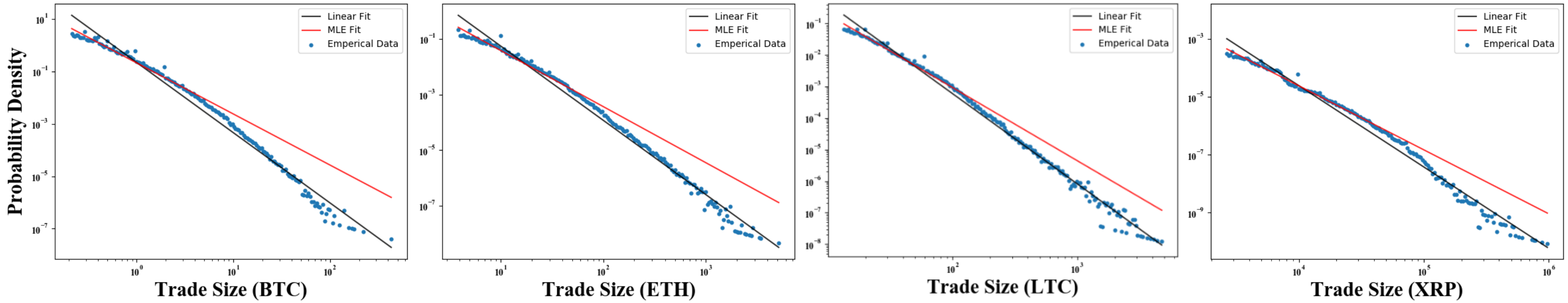
UT6


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

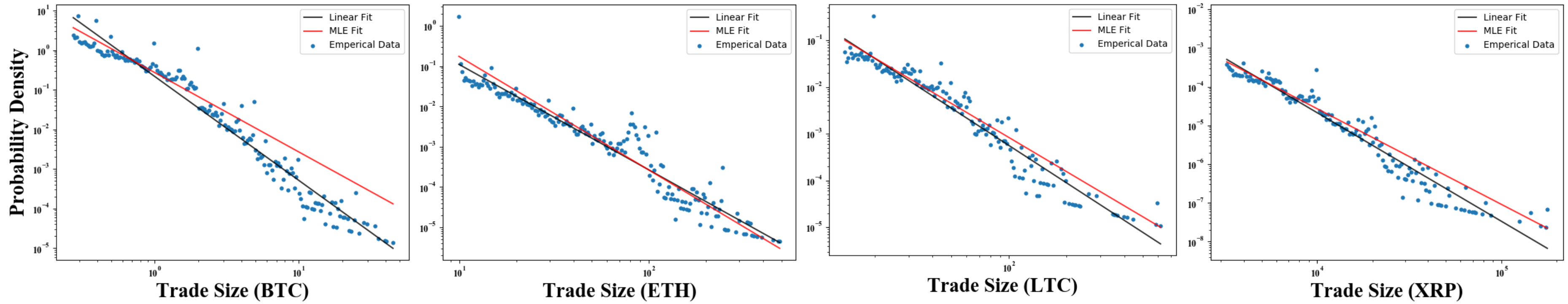
UT7


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (XRP)

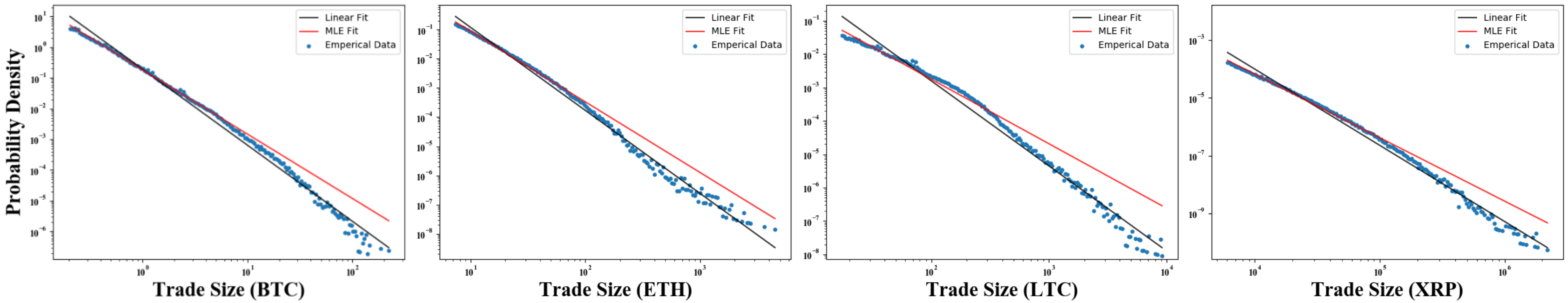
UT8


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

UT9

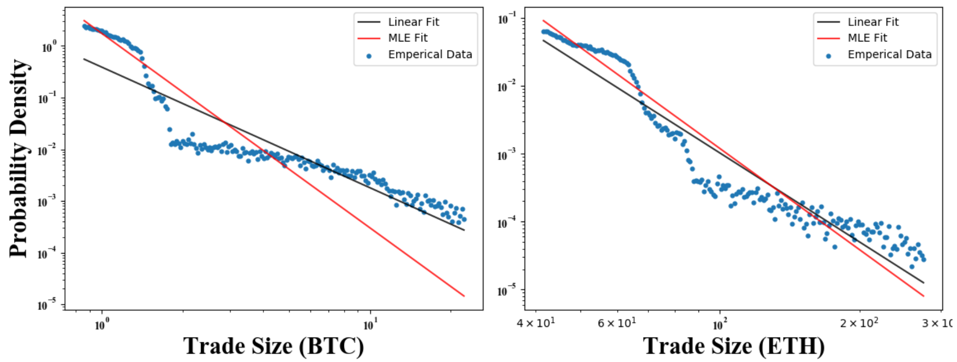

Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

UT10

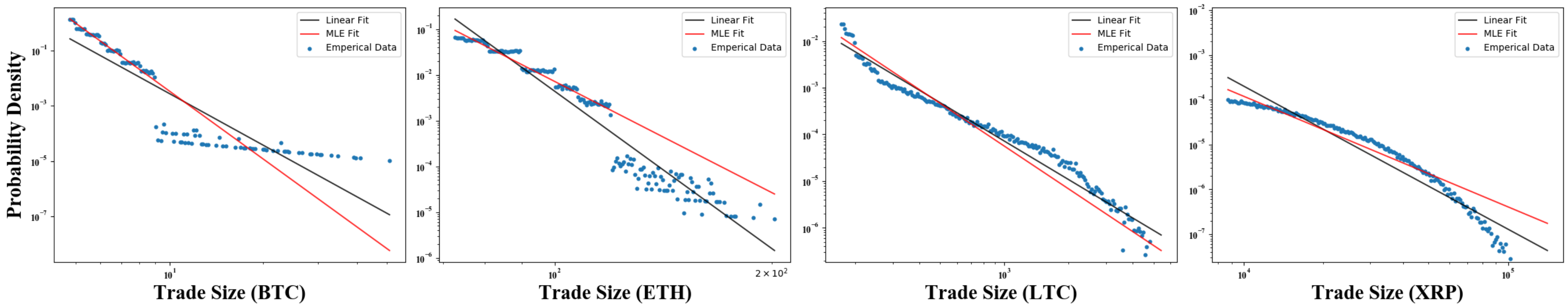


Panel U: Unregulated Tier-2 exchanges

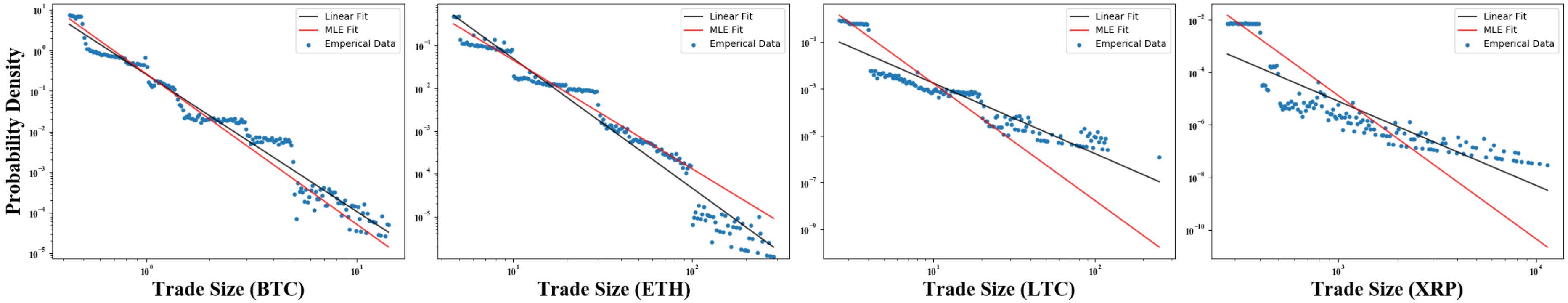
U1



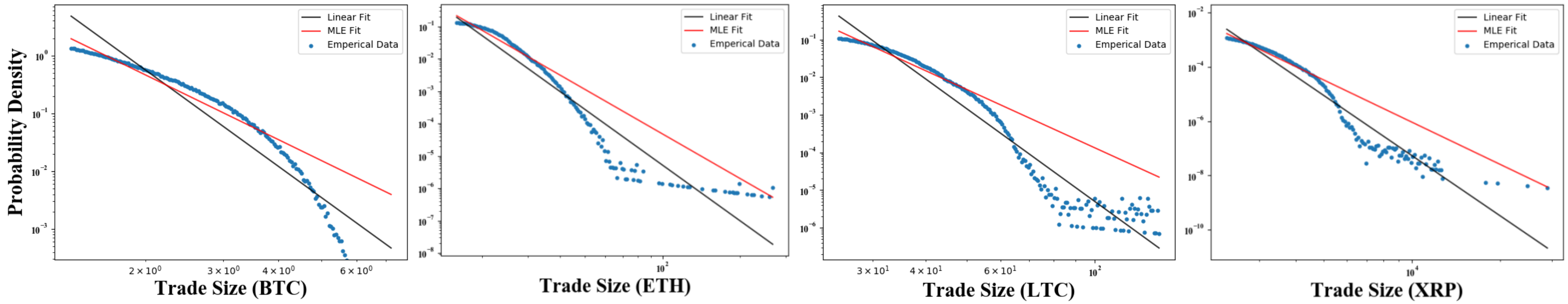
U2



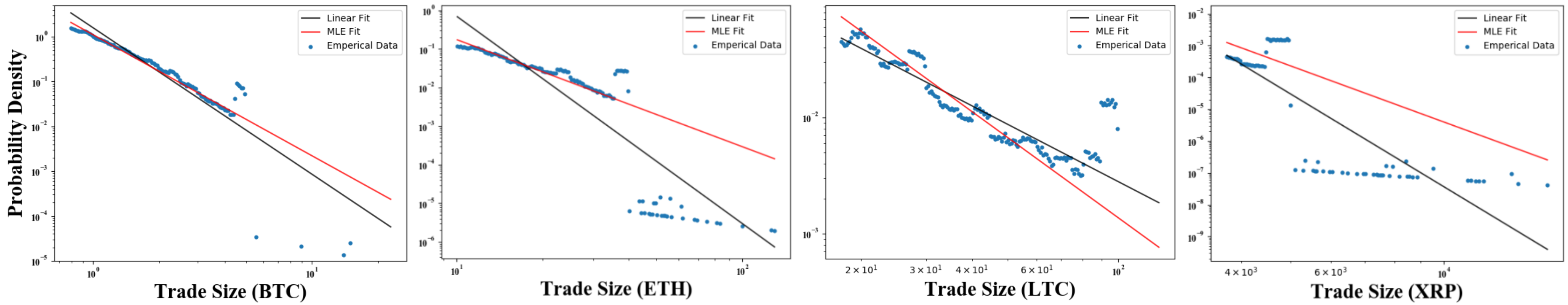
U3


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

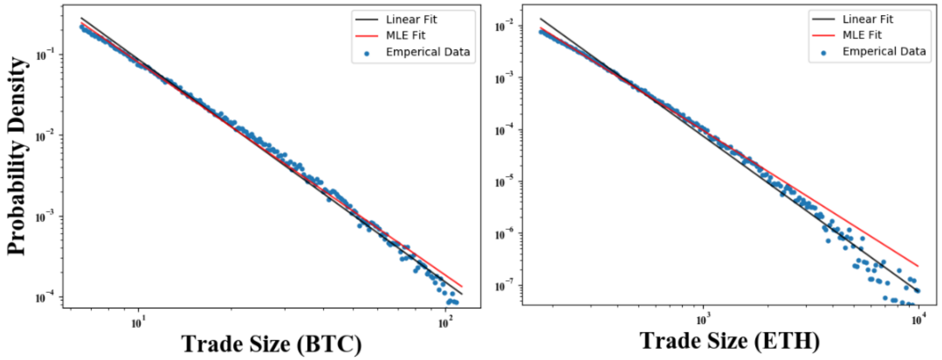
U4


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

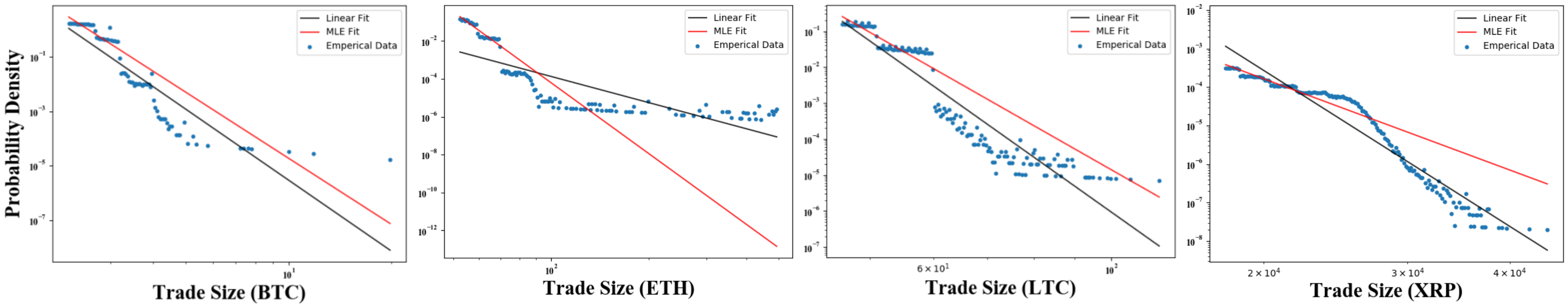
U5


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

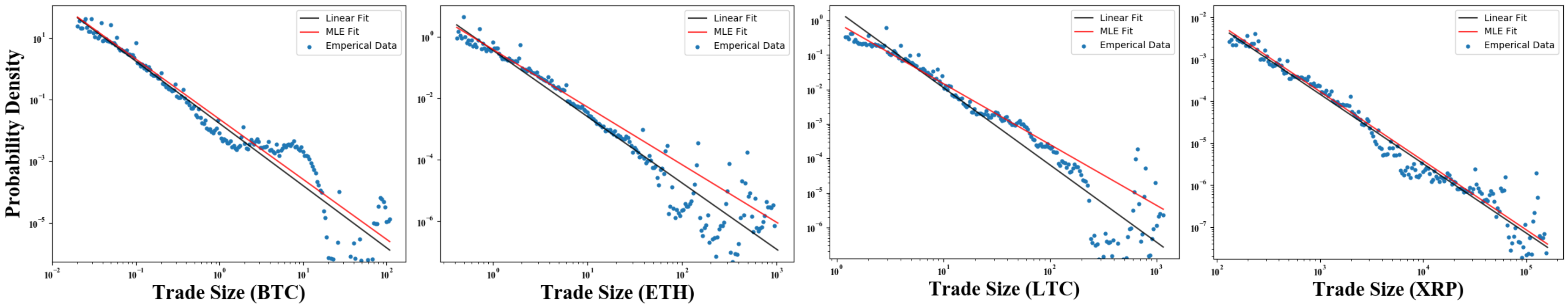
U6


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)

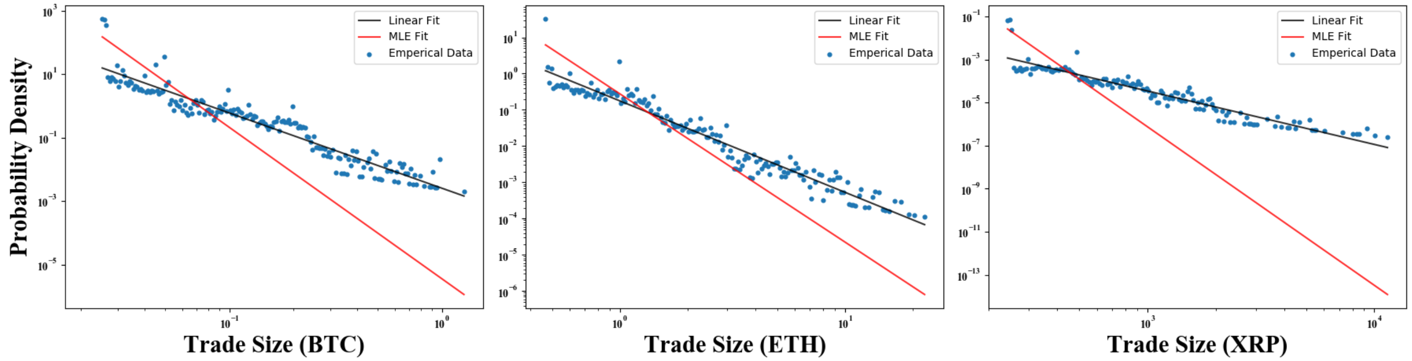
U7


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

U8

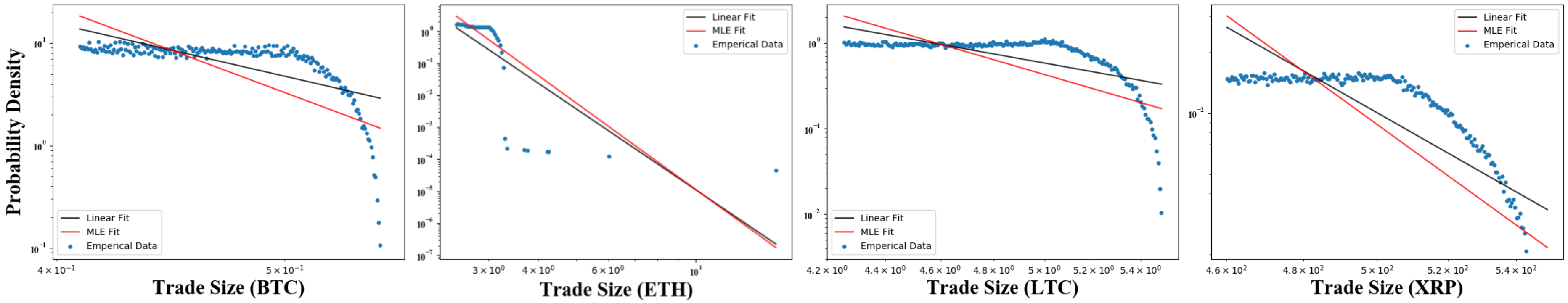

Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

U9



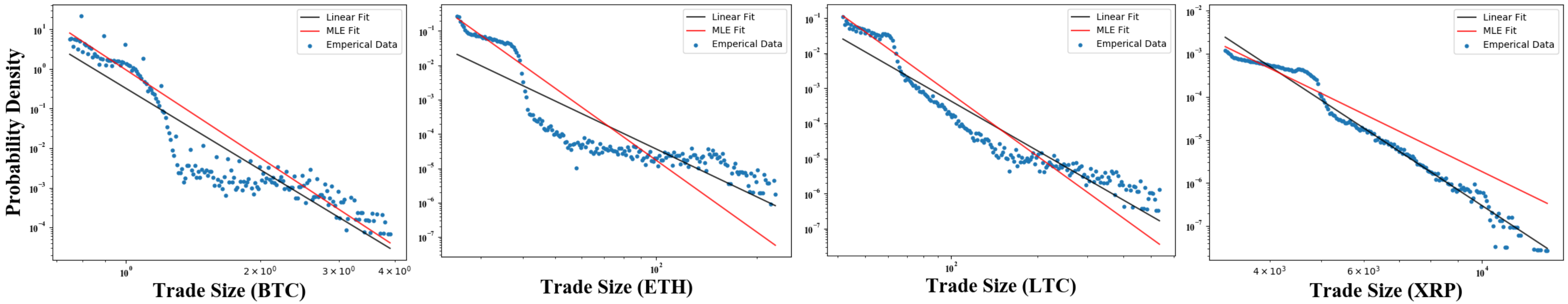
Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

U10



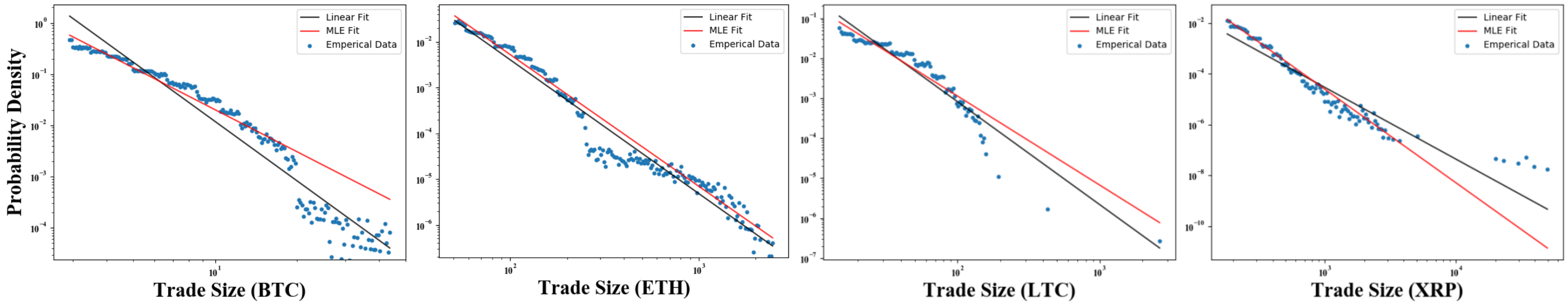
Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

U11



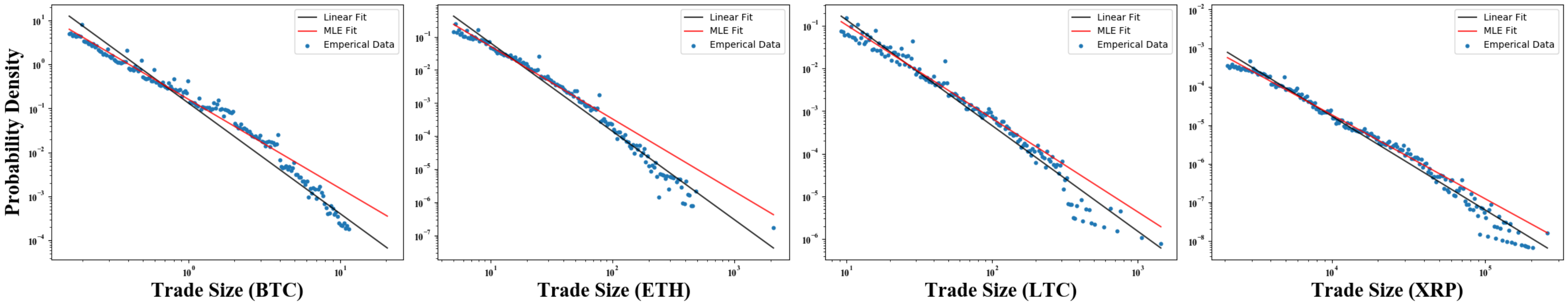
Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (ETH)
Linear Fit
MLE Fit
Emperical Data
Trade Size (LTC)
Linear Fit
MLE Fit
Emperical Data
Trade Size (XRP)

U12



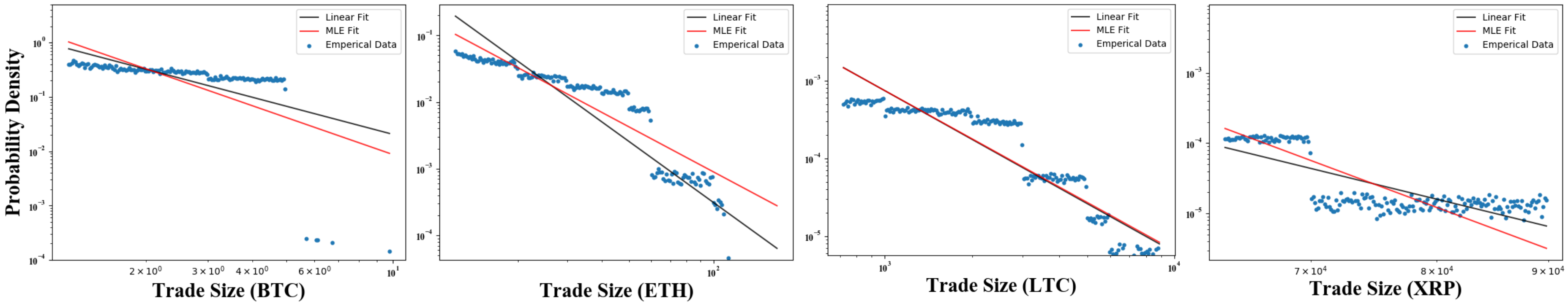
Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U13



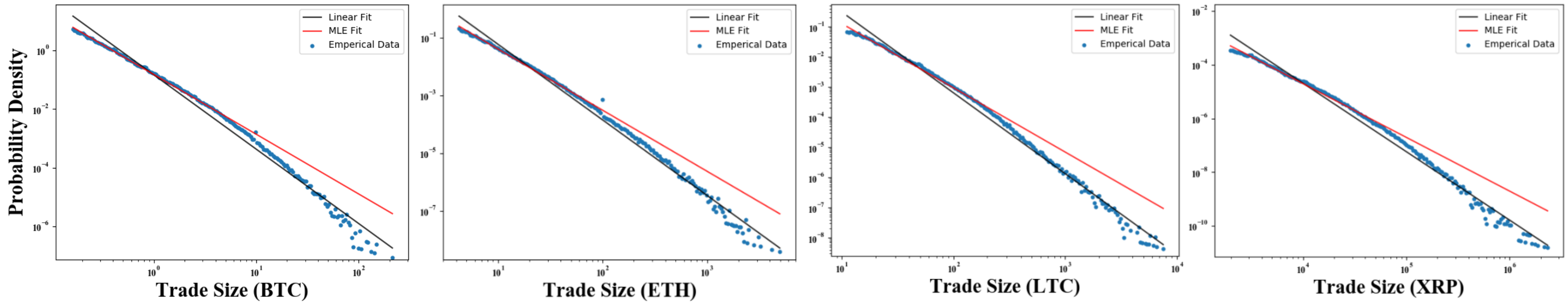
Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U14

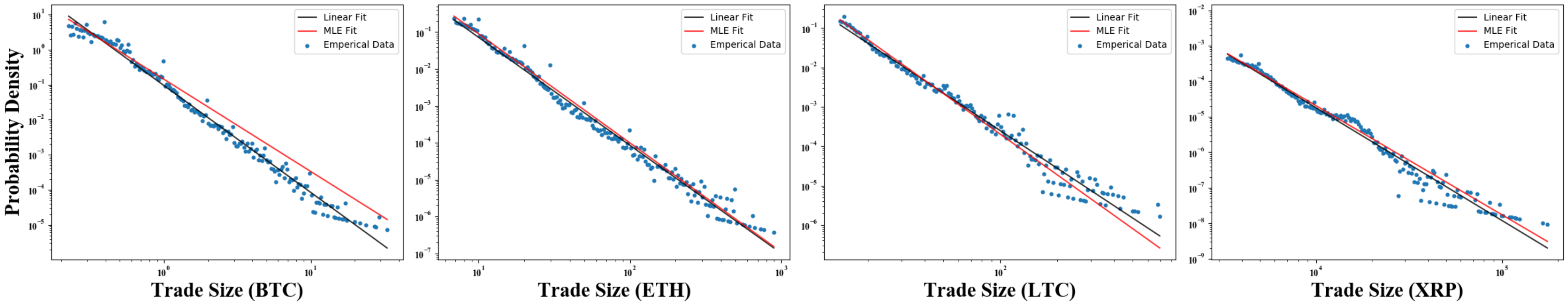


Probability Density
Linear Fit
MLE Fit
Emperical Data
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U15


Linear Fit
MLE Fit
Emperical Data
Probability Density
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

U16


Linear Fit
MLE Fit
Emperical Data
Probability Density
Trade Size (BTC)
Trade Size (ETH)
Trade Size (LTC)
Trade Size (XRP)

Appendix F

Appendix F summarizes statistical tests in Section 4 for each crypto exchange, including Chi-squared tests for Benford s Law, t-test for trade size clustering, scaling exponents for power law. For each test, we report four cryptocurrency pairs, BTC, ETH, LTC and XRP. For each exchange, percentage of failed test is calculated as the number of failed tests over the total number of tests across cryptocurrency. Similarly, percentage of failed tests for each type of cryptocurrency is measured as the number of failed test at 5% significance level over the total number of tests in one type of cryptocurrency. In the table below, value of variable is 1 for failure of test and 0 for pass of test. In Chi-squared tests of first significant digits, "failure of test" is defined that exchange failed to conform to Benford s Law or trade pattern predicted by regulated exchanges, statistically significant at 5% or 1% level; "pass" otherwise. In t-test of size clustering, "failure of test" refers to that exchange do not show apparent size clustering at multiple of 100 units while "pass" represents noticeable clustering effect at 5% or 1% significance level. In estimation of power law exponents, "failure of test" refers to that scaling exponent either $\hat{\alpha}_{OLS}$or $\hat{\alpha}_{Hill}$ lie outside the Pareto–Lévy range (1, 2) and tail distribution does not show linear trend on log-log plot; "pass" otherwise. In Chi-squared tests of roundness level, "failure of test" are defined that unregulated exchange show different roundness distribution from regulated exchanges, at 5% or 1% significance level; "pass" otherwise.

Cryptocurrency	BTC	ETH	LTC	XRP
Number of failed tests	37	33	31	39
Total Number of test	87	87	75	81
Percentage of failed tests	42.53%	37.93%	41.33%	48.15%
Total Number of test(without regulated)	78	78	66	72
Percentage of failed tests(without regulated)	47.44%	42.31%	46.97%	54.17%

Exchange Code	**Benford s Law** Inconsistent with Benford s Law				**Clustering** No apparent clustering at multiples of 100 units				**Power Law** Outside 1<α<2 Nonlinear trend on logarithm scale				**Percentage of failed tests**
	BTC	ETH	LTC	XRP	BTC	ETH	LTC	XRP	BTC	ETH	LTC	XRP	
Panel A Regulated exchanges													
R1	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
R2	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
R3	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
Panel B Unregulated Tier-1 exchanges													
UT1	0	0	0	0	0	0	0	1	0	0	0	0	8.33%
UT2	0	0	0	0	1	0	0	0	0	0	0	0	8.33%
UT3	1	0	0	1	0	0	0	0	0	0	0	0	16.67%
UT4	0	0	0	0	1	1	1	1	1	1	1	1	66.67%
UT5	0	0	0	0	0	0	0	1	0	0	0	0	8.33%
UT6	0	0	0	0	1	1	1	1	1	1	1	1	66.67%
UT7	1	0	NA	0	0	0	NA	0	1	0	NA	0	22.22%
UT8	0	0	0	0	0	0	0	1	1	0	0	0	16.67%
UT9	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
UT10	0	0	0	0	1	1	1	1	0	0	1	0	41.67%

Panel C Unregulated Tier-2 exchanges													
U1	0	0	NA	NA	1	1	NA	NA	1	1	NA	NA	66.67%
U2	1	1	0	0	1	1	1	1	1	1	1	1	83.33%
U3	0	1	0	0	0	0	1	1	1	1	1	1	58.33%
U4	0	0	0	0	1	1	1	1	1	1	1	1	66.67%
U5	1	0	1	1	0	1	1	1	1	1	1	1	83.33%
U6	0	0	NA	NA	1	1	NA	NA	0	0	NA	NA	33.33%
U7	0	1	1	0	1	0	0	1	1	1	1	1	66.67%
U8	1	1	1	1	1	0	1	1	1	1	1	1	91.67%
U9	1	1	NA	1	0	1	NA	1	1	1	NA	1	88.89%
U10	0	0	0	1	1	1	1	1	1	1	1	1	75.00%
U11	0	0	0	0	0	0	0	1	1	1	1	1	41.67%
U12	0	0	0	0	1	1	1	1	1	1	1	1	66.67%
U13	0	0	0	0	0	0	0	0	0	0	0	0	0.00%
U14	0	1	1	1	1	1	1	1	1	1	1	1	91.67%
U15	0	0	0	0	1	1	1	1	0	0	0	0	33.33%
U16	0	0	0	0	0	0	1	1	1	0	1	1	41.67%

Appendix G

Appendix G summarizes Chi-square tests for trade size distribution and level of roundness in Section 4 for each crypto exchange, in which pattern of regulated exchanges is used as benchmark. For each test, we report four cryptocurrency pairs, BTC, ETH, LTC and XRP. In the table below, value of variable is 1 for failure of test and 0 for pass of test. "failure of test" is defined that unregulated exchange failed to conform to trade pattern predicted by regulated exchanges, statistically significant at 5% or 1% level; "pass" otherwise.

Exchange Code	**Trade size distribution** Inconsistent with Regulated exchanges				**Level of Roundness** Inconsistent with Regulated exchanges			
	BTC	**ETH**	**LTC**	**XRP**	**BTC**	**ETH**	**LTC**	**XRP**
Panel A Regulated exchanges								
R1	0	0	0	0	0	0	0	0
R2	0	0	0	0	0	0	0	0
R3	0	0	0	0	0	0	0	0
Panel B Unregulated Tier-1 exchanges								
UT1	1	1	1	1	0	1	1	1
UT2	0	1	1	0	0	0	0	1
UT3	0	0	0	1	1	0	0	1
UT4	1	1	1	1	1	1	0	1
UT5	1	1	1	1	1	0	1	1
UT6	1	1	1	1	0	1	1	1
UT7	1	0	NA	0	1	1	NA	1
UT8	0	1	1	1	0	1	1	1
UT9	0	1	1	1	0	0	1	1
UT10	0	0	1	1	1	1	1	1
Panel C Unregulated Tier-2 exchanges								
U1	1	1	NA	NA	1	1	NA	NA
U2	1	1	1	1	1	1	1	1
U3	1	1	1	1	1	1	1	1
U4	1	1	1	1	1	1	1	1
U5	1	0	1	1	1	1	1	1
U6	1	1	NA	NA	1	1	NA	NA
U7	1	1	1	1	0	1	1	1
U8	1	1	1	1	1	1	1	1
U9	1	1	NA	1	1	1	NA	1
U10	1	1	1	1	1	1	1	1
U11	1	1	1	1	1	1	1	1
U12	1	1	1	1	1	1	1	1
U13	1	1	1	1	1	1	1	1
U14	1	1	1	1	1	1	1	1
U15	1	1	1	1	1	1	1	1
U16	0	0	0	0	1	1	1	1