Estimating imperfections of cryptocurrency markets by calculating potential arbitrage profits

Hugo Joncour
Gina Cody School of Engineering and Computer Science
Concordia University
Canada
hugo.joncour.k2o@gmail.com

Keywords— Cryptocurrency, Trading, Arbitrage, Finance

I. INTRODUCTION

The Bitcoin is a digital asset designed as an alternative to traditional fiat currencies, its purpose is to create a system without relying on a central bank or third party for transactions [1]. The Bitcoin uses the blockchain, a decentralized technology of encryption which allows each transaction to be both public and traceable. Most cryptocurrencies do not share the same ambition of being an alternative to pre-existing currencies and are designed with a more specific purpose in mind [2,3].

Bitcoin was introduced as a means of exchange for daily payments, but its properties allowed it to be used in many other ways, including speculation [4], cross-border money transfer [5], and money laundering [6].

The value of bitcoins, and other cryptocurrencies is largely determined by the trading that takes place on the online *exchange* [7]. Exchanges are the most common way to trade a cryptocurrency, by allowing a client to buy, sell or store his digital assets on the platform. Most of these platforms allow users to put 'buy' and 'sell' orders for a certain price for any currency offered on the exchange, like financial platforms for commodities. On each platform, each currency can be traded with most other cryptocurrencies available on the platform. Each of these trades is agreed on an exchange rate expressed in the cryptocurrency's unit price.

Over the last 5 years the capitalization of cryptocurrencies went from less than 5 billion USD to more than 340 billion today [8]. This explosion of the cryptocurrency market inspired many people and companies to create their own exchanges and cryptocurrencies: as of today, we currently count 7254 cryptocurrencies, 332 recorded platforms, and 30 550 markets (a cryptocurrency X, traded for a cryptocurrency Y, on a specific platform Z).

Even though these trades platforms involve millions of people, that the information is publicly available and can even be automated with APIs, it is easy to observe arbitrage opportunities. When one user can buy a cryptocurrency one exchange, send it to another exchange, and sell it for a higher price. These arbitrages can be performed by anyone with an account on the platform trading these cryptocurrencies.

This document intend to analyze the arbitrage opportunities at a time T, and from this, try to evaluate how big 'market imperfections' are for an environment where information is free and available at any time, where transaction fees are low, can be processed in a short amount of time, and where economic atomicity exists.

From these results, more research could estimate the time necessary for an information (a currency rate) to propagate, and measure how long it takes for a considerable number of homogeneous markets to align with the information, or for an arbitrage to take place and 'absorb' the market's imperfection.

As financial arbitrages are not complex financial transactions (they only require finding somebody selling below market price and somebody else buying above market price), an algorithm efficient at finding arbitrage opportunities (and deciding which exchanges and cryptocurrencies the money flow should take) can be the foundation of a trading strategy, only requiring the programming implementation, and the capital to start trading.

As it is impossible to know at any moment if a new exchange appears, or if a new cryptocurrency is created, I will use information from the website coinmarketcap.com to keep track of the cryptocurrencies and exchanges.

II. METHODOLOGY

This article will evaluate how big market imperfections can be in an environment where almost all economical 'perfect competition' requirements are present, at a time T.

A. Dataset description

My initial dataset includes data from 200 exchanges, 3055 cryptocurrencies and 21340 markets, from October 9th, prices were from 6PM.

The volume traded T of all 7537 cryptocurrencies, on the 349 recorded exchanges in the last 24 hours was estimated at 91 559 435 098 USD. In comparison, the volume traded of my dataset amounts for 78 071 059 706 USD. However, both the different market pairs of cryptocurrencies (Figure 1) and the volume Traded (Figure 2) are unevenly distributed among exchanges.

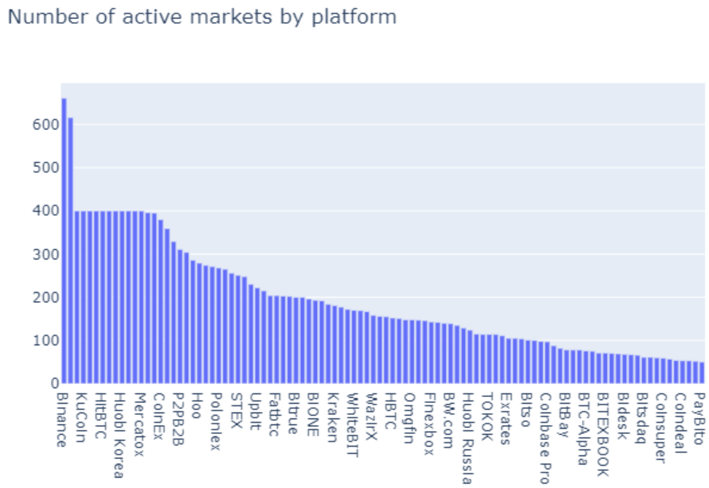

Figure 1: Number of cryptocurrency pair by platform (only the 100 platforms with the most pairs appear on the above figure). Data from Coinmarketcap on October 9th 2020, ploted with Plotly.

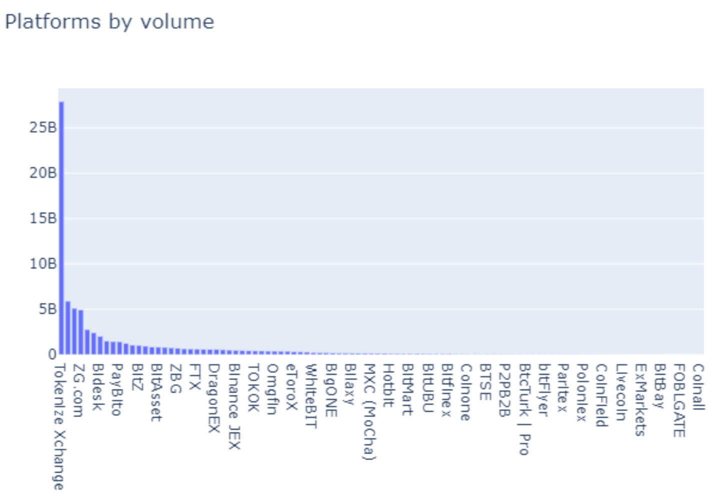

Figure 2: Volume traded by platform (only the 100 exchanges with the highest volume traded are represented on the above figure). Data from Coinmarketcap on October 9th 2020, ploted with Plotly

B. *Graph Theory approach*

This problem can be formulated as a Graph Theory problem where the capital of an investor looks for a path across exchanges where vertices can have a positive or a negative cost. If at the end of the path the capital of the investor is bigger than at the origin, the investor made profit.

a) Modeling the problem

This We can pose our data as a Maximum Flow problem.
Let $N = (V, E)$ be a Network N with $s, t \in$ N where s and t are the source and the sink of N.
The Vertices V will represent the exchanges, and for each cryptocurrency existing on two platform, there will be an edge E, with infinite capacity between the two.
For each exchange E_i on the network, there exist a subnetwork $N_i = (V_i, E_i)$ with $s_i, t_i \in \mathrm{N_I}$ and where s_i and t_i are respectfully the source and the sink of the subnetwork N_i.

In the subnetwork N_i , the Vertices V_{ij} represent the cryptocurrencies traded on the exchange V_i while the Edges E_{ij} represent the markets.
The capacity c_{ij} of an edge E_{ij} is given by:

$$c_{ij} = \frac{V_{ij}}{R_{ij}} \tag{1}$$

where V_{ij} is the volume (in dollar USD) traded that day on this market j and where R_{ij} is the rate of exchange with respect to the unit of the market's pair.

b) Graph Theory Algorithm

The following algorithm, inspired by the Ford-Fulkerson algorithm, gives the total arbitrage and the path with the list of trades to make.
To run, the algorithm needs markets to be ordered by price for each cryptocurrency, its complexity is:

$$O(E_{ij} \cdot \mathrm{V_{ij}} \cdot \frac{\text{Volume}}{\text{capital}}) \tag{2}$$

Algorithm:GraphPaths(capital,Currencies[])

```
Input: capital, Currencies [];
Output: Augmenting paths and total arbitrage
Path = [];
for i in Currencies do
  while Currencies_i.market[0]<Currencies_i.market[n]:
  do
    volume = min(market_i0, market_in, capital);
    market_i0, market_in−=volume;
    if capital>volume then
      capital=capital−volume+volume·(market_in.price=market_i0.price);
    else
      capital+=volume·(market_in.price=market_i0.price);
    end if
    Path.addCurrencies_i.market[0], Currencies_i.market[n];
    if market.volume ≤ 0 in Currencies: then
      remove: .market[x];
    end for
    return Path, capital
```

c) Limitations of the Graph Theory approach

This Though the Graph Theory approach will give a solution that can be verified by giving a result, the max-flow, and it is proof, the list of transactions between each node to obtain the market price on every platform, it is only a theorical approach that assumes:

1) The algorithm finds a path before conditions changes.
2) All transactions are executed.
3) No transaction cost.

1) and 2) are equivalent to assume that the algorithm finds instantly a solution, and that all transactions of the solution are instantly executed, which are unrealistic.

C. Statistical Approach

A statistical approach is another method that can estimate the total of price anomalies across all platforms and all cryptocurrencies.
By linearity of the mean, by computing for each cryptocurrency its average price across all exchanges, and then calculating for each with its adjusted volume we get the total anomaly for that cryptocurrency.

a) Algorithm to Compute

Algorithm: statEstimation(Cryptocurrencies[])

```
Input: capital k, Cryptocurrencies [];
Output: Sum of theoretical arbitrage on those markets
Total = 0;
for i in Cryptocurrencies do
    mean = 0;
    for market in Cryptocurrencies_i do
        mean+= market.Price;
    end for
    mean = mean / markets.size;
    deviation = 0;
    for market in Cryptocurrencies_i do
        deviation+= (|market.Price−mean| / market.Price) · (market.volumeTraded / Time.scale);
    end for
    Total += deviation;
end for
return Total
```

b) Limitations of the Statistical Approach

The main advantage of the statistical approach is that it can be computed in a polynomial amount of time. However, this method does not give the solution, or the proof of the result, as it does not indicate which arbitrages can be done to reach such profit.
This approach assumes the volume traded is globally equally distributed throughout the day.
This algorithm only gives the theoretical profit that can be made at time T.

III. Literature Review and Results

Since the emergence of Bitcoin and exchanges, more and more people took interest in cryptocurrency trading. Many different approaches to trading were adapted to the cryptocurrency's markets: technical analysis [9], statistical analysis [10], machine learning [11], pricing [12], arbitrage [13] and high frequency algorithms [14]. Many of these methods are combined to obtain better results. Arbitrage is a strategy using the temporary inconsistency in the price of an asset. This article compares the research on the use of these methods combined to arbitrage to trade on cryptocurrency exchanges.

A. Current state of arbitrage strategies between cryptocurrencies

The current literature about arbitrage on cryptocurrency markets currently enumerates many different approaches. Some articles seem to focus on the methods used to find and solve price discrepancy [12] or the technical implementation of solutions [13][14], while some others focus on the more theoretical part, the existence of the price dispersion [15][16][17].

The above literature supports the idea that many arbitrage possibilities exist. Some authors base their research on algorithms outperforming current trading algorithms [14], some others demonstrate that arbitrage opportunities are under-exploited by market participants [16].

Most of the arbitrage from the above literature is executed on one platform between a few cryptocurrencies (triangular arbitrage) [18] [19], only two papers applied arbitrage between five and six exchanges simultaneously [10] [14]. On the opposite, three out of four of the articles researching arbitrage market discrepancies used between five and seventeen different platforms.

There does not seem to be any current article about arbitrage using data from more than 17 different exchanges, which seems to be a gap because the research [15] shows that price deviations are much larger across exchanges, especially when both are located in different countries.

Unlike the articles above, the present paper aims at estimating the total amount of price discrepancy at a time T, which is the exact theoretical profit that can be made if that price incoherence was solved. The authors of the cited articles focused more on the 'possible' by testing trading strategies, or by proving that arbitrage is possible, but focused on much smaller samples of data, which don't give a correct idea of how big price discrepancy globally is.

Price anomalies and cryptocurrency markets is a topic of interest for many authors but there is not currently any article aimed at estimating how big can it be with the biggest data set. This article estimates price variations across much more exchanges and for more cryptocurrencies than any published article, using two different methods, a trading approach, and a statistical approach.

B. Results

Results from the statistical method gives a raw estimation of 200 063 807 USD of arbitrage at any time of the day for the set of data with a volume of 78 071 059 706 USD.

Due to the complexity of the problem (finding the optimal path on a graph being a NP-hard problem) the graph theory method couldn't be computed on the whole dataset. By running the algorithm on five random samples of the dataset with a volume of 1 billion USD and extrapolating to the total dataset volume, all results were matching the statistical estimation with a margin of ±10%.

$$\frac{200\,063\,807}{278\,071\,059\,706} \approx \frac{1}{390} \approx 0.00256 \qquad (3)$$

The current dataset estimates that theoretically every minute ~ 0.00256% of the daily volume can be arbitraged, which gives an average of ~ 3,69% of the volume traded is pure potential profit.

IV. SUMMARY, CONCLUSION AND DISCUSSION

A. *Summary*

Using two different methods, this article gives an estimation of how big price anomalies are on cryptocurrency markets at a time T.

The statistical method finds the estimation by calculating how big price gaps were across exchanges for identical trading pairs and comparing their volumes, while the trading approach gives a path through exchanges and pairs and calculating how much profit could be made by buying at the lowest price and selling at the highest price for all available trades.

Both algorithms have advantages and disadvantages but gave matching results which is a sign of reliability of both methods. The estimation results seem correct, with an estimated of 3,69% of theoretical profit on average.

B. *Conclusion and Discussion*

Even though both techniques might estimate correctly with the information given, they rely on assumptions and were used on a unique dataset, and it would be wiser not to assume that the present results of the present article can be extrapolated to any other moment in time.

More knowledge of the price mechanism could be acquired with further research and replicated experiments with large datasets such as the one used. Both algorithms and results could be improved substantially with a more accurate dataset, and observations on more than one day could help conclude on a more solid relation between price variation and arbitrage with respect to volume.

V. REFERENCES

[1] S. Nakamoto, "*Bitcoin: a peer-to-peer electronic cash system"*, October 2008.

In this article published in a mailing list of the forum metzdowd.com, Satoshi Nakamoto, an unknown man, introduces the Bitcoin, a peer-to-peer currency relying on encryption that will allow online payments without any intervention from any financial institution. This paper describes how each transaction includes the previous owner's hashed public key to make prevent the problem of double-spending. Transactions are secured by a proof-of-work algorithm, and as long as the majority of transactions are processed by honest nodes, all transactions are safe. He explains how the network works, and how secured transactions reward miners with new coins, an incentive to support the network. He also addresses privacy as well as vulnerability concerns.

Despite not being published in a peer reviewed journal, this article is one of the most important for the cryptocurrency field and the encryption algorithm that is uses to protect transactions inspired hundreds of other currencies.

This article is used as a source to introduce the Bitcoin because it explains what it is, what it relies on, why it was created and how it works. This article also helps to understand the other cryptocurrencies, that it inspired.

[2] Z. Amdsen, R. Arora, S. Bano, "*The Libra Blockchain"*, July 2019.

This article is the white paper of Facebook's Cryptocurrency's Blockchain. It introduces the technology on which Facebook's cryptocurrency, Libra, will work and what it aims toward. In this white paper, the many authors describe their technology. Unlike the Bitcoin ('permissionless'), to take part to the monetary creation and encryption process, Libra requires 'miners' to disclose their identity. Governance of the technology is assured by the Libra Association and its 28 founding members (partner companies such as Uber, PayPal, MasterCard). The white paper states that the association wants the technology to evolve through time, toward a more decentralized and open process.

This article introduces a unique cryptocurrency that is closer to traditional institutions than other cryptocurrencies -with private corporations taking part in the decision process- and at the same time more innovative than any current traditional currency. Many observants voiced concerns regarding the power it gives to corporations. However, the paper is biased and tries to give a more user-friendly description of the system than it really is by using words such as 'decentralized' and open, even though most of the change in that direction is yet to be done.

This source gives an example of how companies inspired by the Bitcoin's success decided to use the idea of an electronic cash system but without the original idea's ideals of decentralized technology, democratic decision taking, transparency and absence of institution.

[3] B. Chase, E. MacBrough *"Analysis of the XRP Ledger Consensus Protocol*" February 2018.

In this whitepaper published on arxiv.org, two engineers from the Ripple company, Brad Chase and Ethan MacBrough, explain the technology used by their company, Ripple (XRP), a cryptocurrency built on a distributed open source protocol that has specialized in international payment settlement for banks. The article describes the Byzantine fault tolerant protocol used by the XRP, the XRP Ledger Consensus Protocol, a different algorithm than the Bitcoin's proof of work, or the Ethereum's proof of stake. Instead of using the Blockchain protocol like the Bitcoin, the XRP uses a distributed consensus mechanism to validate transactions. The nodes on the Network vote on validity of each transaction, which makes the XRP protocol much faster, and far more energy efficient than other cryptocurrencies relying on the proof of work.

Every day 2 billion of dollars of transactions rely on the algorithm described in this article, but it does not bring much innovation to the cryptocurrency community. The governance of the XRP is done by companies and are less transparent in their decisions than other committees. The XRP is the biggest cryptocurrency with the most centralized decision taking and requires trusting a company rather than the community.

This white paper is a source because it is the most important example of a company using a general-purpose cryptocurrency with business specific applications.

[4] E-T Cheah, J. Fry, "*Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin", Elsevier, February 2015.*

In this article, Eng-Tuck Cheah, associate professor at Nottingham Business School, and John Fry, senior lecturer at University of Bradford, challenge the idea that bitcoin is a valid currency as it is not recognized as such by institutions and because it does not fulfill the fiat currencies functions: unit of account, means of payments and store of value. The authors then proceed to use statistical methods to test for the probability of a bubble using the data from Coindex, from 2010 to 2014. One of the observations is that Bitcoin seems to be more an object of speculation than a functional money and that its value in the long term is 0. After applying more tests to the data, the authors conclude to a bubble, that it accounts for almost half of the observed prices, and reject the null hypothesis in favor of a nonlinear superexponential bubble.

This article lacks credibility because the Bitcoin's evolution has since contradicted their predictions. It seems that the authors have an anti-Bitcoin bias and treat it like they would treat a regular asset, which have very different properties than the Bitcoin.

This source will be used to prove that an important part of the Bitcoin's price, and by extension, many of the cryptocurrencies price are affected by bubbles, and that their price is driven by speculators.

[5] V. Iaralov, I. Gordon, D. Woo, "Bitcoin: a first assessment", Merrill Lynch, p. 7, December 2013.

In this report, David Woo, Ian Gordon and Vadim Iaralov, currency experts, try to assess the Bitcoin's potential as a digital currency and as a speculative asset for Bank of America: Merrill Lynch's investors. The report gives a good description of how Bitcoins are produced and then explains how the sector is witnessing a huge growth with important investments being made in Bitcoin start-ups and exchanges. The report tries to assess the Bitcoin's 'fair value', as a medium of exchange, and as a store of value: they proceed to evaluate how much it's maximum worth can be in both case and compare its potential competitors in each sector. The article also includes a cost-benefit analysis that is still valid today.

Although this assessment was published in December 2013, a year during which the Bitcoin's price increased 100 fold and despite underestimating the price it would reach in the following years, this report is quite accurate and most of the potential they saw in the Bitcoin was right.

This source is included to show that as early as December 2013 the Bitcoin was used to transfer money through borders at low cost.

[6] J-J. Oerleman, O. Van Deventer, R. Van Wegberg, "*Bitcoin money laundering: mixed results"*, Journal of Financial Crime, May 2018.

Rolf van Wegberg, Jan-Jaap Oerlemans and Oskar van Deventer work respectively for the Technical University of Delft, The Ministry of Security & Justice, and the TNO research institute. In this article they investigate how cybercrime uses Bitcoin laundering schemes to wipe incriminatory money trail. The article investigates how some online services offer to anonymize bitcoins further by dissociating the coins on a wallet from the transactions they were used in to cover the tracks linked to criminal activities. Bitcoin are associated to a wallet, through an address, but not an individual, unlike bank accounts. Since all transactions are registered in the blockchain, it is possible to look through which addresses money has traveled before arriving on a bank account. The authors tested five different '*Bitcoin Mixers*', three out of five turned out to be scams (the money was stolen and never reached the destination wallet), while the last two performed successfully and no trail could link the money to the initial wallet before the mixing.

This article is the first one to investigate this kind of activities. Some part of the investigation is now outdated (where the authors compare which platform provides the most anonymity when trying to exchange bitcoin for fiat currencies) because PayPal does not accept bitcoins anymore, but this article is still extremely relevant for anything that involves bitcoin and online criminal activities.

This source will be used to demonstrate that Bitcoin is used un laundering money activities, and that Bitcoin itself can be laundered to cover track of criminal activities at a very low fee.

[7] N. D. Bhaskar, D. L. Kuo Chuen, "*Chapter 28: Bitcoin Exchanges"*, in "*Handbook of Digital* Currency", p. 559-560.

[8] Data from coinmarketcap.com;

[9] A. Detzel, H. Liu, J. Strauss, "*Bitcoin: Predictability and Profitability via Technical Analysis*", September 2018. Paris-December Conference 2018.

[10] Deinert, T. Günther Fischer, C. Krauss, "Statistical Arbitrage in Cryptocurrency Markets", "Journal of Risk and Financial Management", February 2019.

[11] S. Leea, N. El Meslmanib, L. N. Switzerc, "Pricing Efficiency and Arbitrage in the Bitcoin Spot and Futures Markets", volume N°53 of "Research in International Business and Finance", October 2020.

[12] J. Lin, Y. Zhou, J. Sun, "Using machine learning for cryptocurrency trading", August 2019, 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS).

[13] C. Păunan "Arbitrage Trading Systems for Cryptocurrencies. Design Principes and Server Architecture", volume n°22 of "Informatica Economică", February 2018.

[14] A. Vo, C. Yost-Bremm, "A High Frequency Algorithmic Trading Strategy for Cryptocurrency", "Journal of Computer Information Systems", December 2018.

[15] I. Makarov, A. Schoar, "Trading and arbitrage in Cryptocurrecy Markets", "Journal of Financial Economics", January 2019.

[16] N. Hajare, "Pricing and Arbitrage in Cryptocurrency Markets", Master"s Tbesis, submitted to Department of Engineering and Electrical Engineering and Computer Science of Massachussets Institute of technology, 2018, accessed on October 28th, 2020.

[17] K. P. Tsang, Z. Yangy, "Price Dispersion in Bitcoin Exchanges", published on "Social Sciences Research Network", May 2020.

[18] Sanghyun Bai, Fred Robinson, "Automated Triangular Arbitrage: A Trading Algorithm for Foreign Exchange on a Cryptocurrency Market", Degree Thesis Project in Technology, Submitted to Electrical Engineering and Computer Science of the KTH Royal Institute of Technology, Stockholm, 2019.

[19] Eric Han, "Intra-exchange Cryptocurrency Arbitrage Bot", Master's Thesis, submitted to the faculty of the Department of Computer Science of San José University, 2018, accessed on October 28th, 2020.