SCHOOL OF BUSINESS ADMINISTRATION SCIENCES

FACULTY OF ACCOUNTING AND FINANCE

POSTGRADUATE STUDIES PROGRAMME ON TAXATION ACCOUNTING AND FINANCIAL MANAGEMENT

Dissertation

ON-CHAIN METRICS ANALYSIS ON THE ETHEREUM BLOCKCHAIN

by

ATHANASIOS GKOURAMANIS

Supervising Professor: Achilleas Zapranis

Submitted as a requirement for the award of the Master's Degree in

Accounting Taxation and Financial Management

September 2023

ABSTRACT

Traditional finance involves two main approaches when it comes to asset valuation, sentiment analysis or other attempted estimations of market variables. The more academically acceptable methods of trying to forecast future magnitudes of financial assets or asset classes based on fundamentals and the attempt to predict price movements using technical analysis, which includes various methodologies.

The recent advent of cryptocurrency and the different ecosystems that came along with this invention, have shown an empty space in traditional financial analysis propositions. Other than technical analysis, which needs nothing more than the historicity of price movements, crypto investors, traders and market analysts usually had no fundamental financial subjects of analysis, since traditional finance and accounting categories were more often than not, non-applicable. This space is being slowly filled with new approaches, which are trying to spot those categories that are essentially fundamental to the formation of crypto asset prices and use them to adequately strategize over future volatility.

The purpose of this paper is to showcase some of the fundamental blockchain data and the subsequent metrics that are used in these types of analyses, as well as examine some of these attempts. Focus is mainly put in the case of Ethereum and its native cryptocurrency Ether because of its nature as a network, on top of which, multiple parallel applications have formed, especially in the decentralized financial sector. Through the use of real examples from market analysts and academic researchers, we try to present an adequate understanding of these new methods and how exactly they are put to use. Finally, we notice similarities and differences with traditional asset classes and traditional methods of analyses, in an attempt to bridge the gap among newer and older approaches to finance and financial analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF GRAPHS

1. INTRODUCTION

The advent of blockchain technology has revolutionised a number of aspects of human activity, from web development to art. The first sector however to be taken aback by it was undeniably finance. As is common knowledge nowadays among people who are familiar with finance and economics, it all started with Satoshi Nakamoto and the first blockchain database, which gave birth to the creation of Bitcoin and changed the world as we know it.

Although, Satoshi's aim with Bitcoin was to create an alternate form of currency that would eventually replace fiat currencies at all levels of transactions on a global scale, something which has not yet occurred and many would argue that it might never come to be; it has unintentionally created a vast array of new possibilities, which have impacted and continue to impact the world heavily in previously unimaginable ways, as is often the case with such inventions.

1.1. Ethereum And The ERC-20 Ecosystem

The largest and most important "side effect" of Bitcoin's invention, so far, has been Ethereum and its subsequent ecosystem. The idea was coined in 2014 by Vitalik Buterin and the network went live in 30 July 2015. At the time, blockchain technology was still unheard of by the majority of people and those who had heard of it, rushed to attribute to Ethereum the characteristic of Bitcoin's competitor, something which could not have been farther from the truth.

The Ethereum blockchain hosts Ether as its native cryptocurrency coin while Ethereum in itself is not a currency but a network, in which Ether is simply used to execute transactions. Many applications can be built on top of this network and can serve a variety of functions from gaming to decentralised finance (known as DeFi). These applications use a particular form of code, called "smart contracts" and each have their own forms of cryptocurrencies, called tokens. The standard protocol, in which these tokens and smart contracts operate in the Ethereum network is called the ERC-20 standard and stands for "Ethereum Request for Comments 20".

1.2. Traditional Valuations Of Financial Assets

The desire to appropriately estimate the true value of financial assets is as long as financial markets themselves. Traditional approaches have looked to investigate a company's or asset's underlying fundamentals, such as balance sheets, income statements, cash flow statements or other accounting related information. Many a financial analyst will examine this and other data alongside key financial indices or performance indicators (for example: Price to Earnings, Revenue Growth, Return to Equity etc.) and employ various criteria to determine whether an asset is overpriced, underpriced or at fair value. Naturally, investors will then opt to add underpriced assets to their portfolios and get rid of overpriced ones, of course considering other factors as well to maximize the efficiency of diversification.

1.3. Technical Analysis

Another very common approach, commonly considered as the antithesis to fundamental analysis, is technical analysis. Technical analysis commonly disregards important economic and financial factors for the determination of the price action of a given asset. Instead, the only variable to be examined is past price action. The technical analysts will make use of purely statistical tools and sometimes plain pattern recognition to try and calculate which direction the price is more likely to move towards. Especially, in the area of cryptocurrencies, traditional methods of fundamental analysis are not particularly popular and in many cases are plainly unusable. A token, as much as its value can be considered to depict the value of the application, of which it is a part, often cannot be connected to accounting categories and a lot of the common financial indices are unable to be constructed. Hence, technical analysis has been, for the most part, the only way to try and estimate the future movement of asset values in the cryptoshpere in the early years of this new and dynamic asset class.

1.4. Financial Analysis On The Cryptosphere Using On-Chain Metrics

As the space of cryptocurrencies and blockchain technology in general is a brand new field, full of potential but also challenges, it doesn't come as a surprise that new methods of analysis have come forward to fill in the blanks of the more traditional ones. A lot of these new methods employ so-called on-chain metrics and they are the ones, which have gathered our interest for this dissertation. Although it will be explained at a greater detail in the coming units, shortly put, on-chain metrics refer to a variety of data, of economic, computing and IT or other nature, which is derived from the blockchain's usage.

This data can later be assessed in order to arrive at conclusions as to what can be expected for a different set of variables in the blockchain. Ultimately, by incorporating it into their research, financial analysts or investors can be aided in determining the likeliest trajectory of an asset's value or be provided with insight to market sentiment or other important factors and make relevant decisions. It's a school of thought regarding analysis, that is quickly gaining traction in the space of crypto investing and DeFi and in the latter pages of this paper, we're going to delve deeper into it and try to create a comprehensive summary of these methods, understand how they work and how people or entities with agency in the crypto markets can make use of them to improve their strategies.

2. A BRIEF REVIEW OF THE EXISTING LITERATURE

The topic of cryptocurrency has been largely overlooked by the academic community in the prior years. However, after 2020 interest has been slowly building up and more papers and dissertations have been coming up in an attempt to dissect and explore this brand new financial and technological ecosystem and make sense of its implications for the totality of economic activity and everyday life.

One such case is the paper by (Urquhart, A.; 2022) in which the author discusses the nature of the Ethereum network and explains its basic functions. The dissertation itself is quite simple in its description of the network and its properties but it goes to show the lack of elemental understanding of the way crypto ecosystems and currencies work, which was prevalent in the wider academic community in the near past. Other studies have attempted to encapsulate the network's underlying programming and unpack valuable insights about blockchain technology and explore alternative future uses, such as (Said, A., Janjua, M. U., Hassan, S. U., Muzammal, Z., Saleem, T., Thaipisutikul, T., Tuarob, S., & Nawaz, R.; 2021).

There are others, such as (Gunay, S., & Kaskaloglu, K.; 2022), who explore the financial connections between Ethereum and the newest addition to the crypto universe, that of NFTs, while (Samreen, N. F., & Alalfi, M. H.; 2023) have recently published a research into decentralized applications and their technical details.

On the topic of valuation and price prediction, significantly more researches have begun to emerge in recent years. (Akgul, A., Şahin, E. E., & Şenol, F. Y.; 2022) examine the use of chaos theory, on-chain analysis and sentiment analysis to determine the valuation of Bitcoin and come to believe that those recent alternatives have the potential to perform better than traditional asset price estimation methods. (Stober, A., & Sandner, P.; 2020) have focused more specifically on market and on-chain metrics, such as monthly active addresses as a proxy for an ERC-20 token's user base and the price of Bitcoin as a proxy for the overall market, applying them to the prices of said tokens and trying to find meaningful correlations. Additionally, (Bakhtiar, T., Xiaojun, L., & Adelopo, I.; 2023) and (Jagannath, N., Barbulescu, T., Sallam, K. M., Elgendi, I., McGrath, B., Jamalipour, A., Abdel-Basset, M., & Munasinghe, K.; 2021) employ quantitative methods, incorporating fundamental factors for the determination of cryptocurrencies' value. Particularly, the former use panel regression models and moving-window regression tests to measure the impact of fundamentals alongside sentiment estimations, like the fear-greed index, on price data of 42 cryptocurrencies while the latter employ variations of an LSTM-RNN model to predict the price of Ether and compare it with more standard versions of LSTM regarding its accuracy. (Kharysh, O.; 2022) also examined neural network models like LSTM and GRU alongside tree based machine learning models like Decision Tree, Random Forest and others, in order to compare and contrast them in their predictive power for the prices of BTC, ETH, and BNB.

Concluding, there is a vast variety of researches and dissertations being made in the whole array of the cryptoshpere and the possibilities for future research seem endless at the moment. However, there is a notable lack of solid academic consensus in each field of study, which leaves researchers with not much to base their work upon. Noting this issue, our paper aims to summarize and put in perspective various different approaches to the field of on-chain data financial analysis, thus setting the field for scientists to have a more focused view while making their initial attempts at grasping and furthering this topic.

3. IMPORTANT DEFINITIONS REGARDING CRYPTO AND THE ETHEREUM ECOSYSTEM

3.1. Essential Crypto Information And Terminology

Although, covering everything crypto related starting from zero is redundant, since the focus of this paper is the financial aspect of the analysis of on-chain metrics, there are some things which need clarification in order to go into the process of investigating metrics, which are produced from the blockchain's function and usage.

To begin with, let's consider the definition of a **blockchain**. A blockchain is essentially a distributed database that maintains an ever-expanding list of organized records, referred to as blocks. These blocks are interconnected through cryptographic techniques. Within each block, there is a cryptographic hash of the preceding block, a timestamp, and transaction data. A blockchain serves as a decentralized, widely distributed, and publicly accessible digital ledger designed to record transactions across numerous computers. This structure ensures that the recorded information cannot be retroactively modified without altering all subsequent blocks and gaining consensus across the network.

This digital ledger comprises scripts, which are programs responsible for performing typical database tasks, such as inputting, accessing, and preserving information. What makes a blockchain distinct is its distributed nature, meaning multiple copies of it are stored on numerous machines, all of which must match for the blockchain to be considered valid. As transactions occur, the blockchain accumulates this transaction data and adds it to a block, akin to a cell in a spreadsheet containing information. Once a block reaches its capacity, the information undergoes encryption through an algorithm, resulting in the creation of a hexadecimal number known as the **hash**. Subsequently, this hash is included in the header of the following block and combined with other data within that block, effectively linking these blocks together in a chain-like fashion.

It's often said that the core characteristic of cryptocurrencies and blockchain technology is the **decentralization** effect. A blockchain introduces a mechanism where data within a database is distributed across numerous network nodes, which can be computers or devices running blockchain software, positioned at various locations. This distribution serves a dual purpose: it not only establishes redundancy but also upholds the integrity of the data. For instance, in the event of an attempt to modify a record within one instance of the database, the presence of other nodes in the network acts as a safeguard, preventing any unauthorized alterations. Consequently, no single node within the network possesses the capability to manipulate the information it holds.

This decentralized distribution, combined with the cryptographic evidence of work completed, engenders a state of irreversibility for the stored information and its historical records, such as transactional data in the context of cryptocurrencies. While a blockchain can certainly house lists of transactions, as seen in the realm of cryptocurrencies, it is also versatile enough to accommodate an array of other data types, including legal contracts, government-issued identifications, or a company's inventory records.

The procedure for transactions varies, contingent upon the blockchain in question. To illustrate, when someone initiates a transaction via a cryptocurrency wallet, which acts as the interface to the blockchain, a predefined sequence of actions is triggered. In the case of Bitcoin, the transaction gets dispatched to a memory pool, where it is temporarily stored and queued for processing until a **miner** selects it for inclusion. When said transaction is incorporated into a block, and that block reaches its capacity with various transactions, it is subsequently sealed and safeguarded through encryption via an encryption algorithm. Following this step, the process of mining commences.

The entire network operates concurrently, engaging in the process of "solving" the hash. In this pursuit, each participant generates a hash that is essentially random, except for a component known as the "nonce," an abbreviation for "number used once." Initially, every miner commences with a nonce value of zero, which is fused with their randomly-produced hash. If the resulting number is not equal to or less than the predetermined target hash, an increment of one is added to the nonce, leading to the generation of a new block hash. This iterative process persists until a miner successfully produces a valid hash, thereby winning the competition and receiving the associated reward. This type of network operation is commonly referred to as “**proof-of-work**” (PoW) because it’s a consensus mechanism, which proves that a certain miner did the work and rewards them in an analogous manner.

Proof-of-work systems were introduced with Bitcoin and remained the only type of consensus mechanism in crypto until recently. Some blockchain networks, like Ethereum, introduced a new system, which is referred to as “**proof-of-stake**” (PoS). Proof-of-stake streamlines the process of block and transaction verification, considerably reducing the computational workload. In PoW systems, the blockchain's security relies heavily on demanding computational tasks. PoS, however, transforms the verification of blocks by harnessing the computing power of coin owners, thus diminishing the necessity for extensive computational effort. Coin owners offer their coins as collateral, a process referred to as "**staking**," in exchange for the opportunity to validate blocks and accrue rewards.

Validators are chosen at random to confirm transactions and validate block data, contrasting the competitive, rewards-based approach of PoW. To become a validator, a coin owner must commit a specific quantity of coins. For instance, Ethereum mandates that a user stakes 32 ETH to operate a node. Multiple validators collaborate to validate blocks, and once a predetermined number of validators validate the block's accuracy, it is then officially confirmed and sealed. Various PoS mechanisms might employ diverse consensus methods. In the case of Ethereum's future introduction of “sharding”, a validator will oversee transactions and incorporate them into a shard block, requiring no more than 128 validators to constitute a voting "committee." Once shards are validated and a block is formed, consensus is reached when two-thirds of the validators concur that the transaction is valid, at which point the block is finalized and sealed.

Both consensus mechanisms serve the crucial functions of synchronizing blockchain data, validating information, and facilitating transactions. Each method has demonstrated its effectiveness in maintaining blockchain integrity, although they come with their respective advantages and drawbacks. Ultimately though, they diverge significantly in their approaches.

In the realm of PoS, individuals responsible for creating blocks are referred to as validators. Validators undertake tasks such as scrutinizing transactions, confirming the validity of activities, participating in voting processes, and maintaining records. Conversely, in PoW, block creators are called miners. Miners engage in solving cryptographic puzzles or hashes to authenticate transactions, and as a reward for this computational effort, they receive cryptocurrency coins.

To assume the role of a block creator, one must "buy into" the position by possessing a substantial number of coins or tokens, which qualifies them as validators in PoS blockchains. In contrast, PoW miners need to invest in specialized processing equipment and endure significant energy expenses to power the machinery engaged in solving computational problems.

The capital-intensive nature of acquiring equipment and managing energy costs in PoW systems results in an elevated entry barrier, thus bolstering the security of the blockchain. In contrast, PoS blockchains reduce the demand for extensive processing power in the validation of block data and transactions. This mechanism also alleviates network congestion and eliminates the rewards-based incentive system that characterizes PoW blockchains.

Table 1: Key differences between PoS and PoW networks (source: investopedia.com)

Proof of Stake	**Proof of Work**
Block creators are called validators	Block creators are called miners
Participants must own coins or tokens to become a validator	Participants must buy equipment and energy to become a miner
Energy efficient	Not energy efficient
Security through community control	Robust security due to expensive upfront requirement
Validators receive transactions fees as rewards	Miners receive block rewards

Having been mentioned briefly in the introductory comments, cryptocurrencies are generally divided in two main categories, a distinction which is fairly important. Cryptocurrencies like Bitcoin and Ethereum are autonomous digital currencies, each operating on its blockchain network and currencies like these are called **cryptocurrency coins**. They primarily function as mediums of exchange, stores of value, or units of account within their respective blockchain systems. The security of these coins is upheld through mechanisms like mining or staking, integral to their operation.

On the other hand, **cryptocurrency tokens** are generated within existing blockchains, such as Ethereum, using smart contracts. Tokens exhibit remarkable versatility, representing a wide array of assets beyond mere currency. They can embody ownership in companies, provide access to specific features within decentralized applications (DApps) or even stand for tangible assets like real estate. Tokens do not possess their independent blockchains; instead, they harness the infrastructure and security of the parent blockchain, granting them adaptability for various applications.

Smart contracts enable tokens to be finely tailored to particular purposes, offering unique functionalities as necessitated by the intended use case. This adaptability, coupled with the capacity to integrate seamlessly with established blockchain platforms, positions tokens as fundamental components in blockchain ecosystems. They power diverse applications, ranging from decentralized finance (DeFi) protocols to digital collectibles in the gaming industry.

Another term, with notable importance in the function of transactions within a blockchain and influence in the economic development of ecosystems is the notion of “**gas**”. This concept represents the computational resources required to perform actions and execute smart contracts on blockchain networks. In Ethereum, users pay **gas fees**, which are priced in tiny fractions of Ether (ETH)—denominations called **gwei** (10^{-9} ETH). Gas is used to pay miners for the resources needed to conduct transactions (in the PoS chain the main gas fee gets “burned” and only the “tips” are paid to the validators). The "**gas limit**" represents the maximum workload a user expects a miner to undertake for a specific transaction. A higher gas limit typically indicates the user's anticipation of a more resource-intensive transaction. Meanwhile, the "**gas price**" refers to the cost per unit of work performed. Consequently, the total cost of a transaction is calculated by multiplying the gas limit by the gas price. In addition to these costs, many transactions may also incorporate "tips," which are added to the gas price. The more substantial the tip, the faster the transaction is likely to be processed. Users who set lower gas limits for their transactions will find themselves lower in the processing queue, impacting the priority of their transactions. It's worth noting that gas prices are subject to the principles of supply and demand. In times of network congestion, gas prices may surge, while they could remain lower during periods of reduced network activity.

3.2.Ethereum

3.2.1. The Blockchain

Having delved into the subject of blockchains and how they operate, we have to go a bit deeper into the particulars of Ethereum, which is the network of our focus and note the important factors in its operation.

The original blockchain, as proposed in Bitcoin, involved simple transactions that transfer some coins from one end-user (typically Alice) to another end-user (typically Bob). The original Bitcoin blockchain can be easily modelled as an abstract data type representing a linked list of blocks of transactions. The accessed data is the cryptocurrency, Bitcoins, and transactions transfer part of the remaining unused assets of Alice to Bob, while keeping the rest with Alice (hence the term Unspent Transaction Output, UTXO to refer to the assets belonging to a client in Bitcoin) (Zakhary, V., Agrawal, D., & Abbadi, A. el.; 2019).

In its inception, Ethereum was also a PoW network and operated in a very similar manner to that of Bitcoin. The launch of the Beacon Chain in December 1, 2020 signalled Ethereum's transition to PoS. The London upgrade introduced EIP-1559, which transformed the network's transaction fee market by introducing a mechanism that dynamically contracts and expands block sizes to deal with transient congestion. For more than 22 months the two chains were running in parallel until finally, on September 15, 2022 the "Merge" became reality with the Paris upgrade. The PoS Beacon chain was merged to Ethereum's Mainnnet, switching off the PoW mining algorithm and switching on the PoS consensus mechanism instead. According to Ethereum's official website, the Merge reduced Ethereum's energy consumption by ~99.95%. The latest important upgrade to the network came in April (Shanghai) and brought along with it a notable increase in the amount of ETH that has been deposited to the ETH 2.0 contract. It is now estimated at a little over 25.6 million, a figure worth some $ 48 billion, with the total percentage increase since the start of 2023 calculated at about 70%. More specifically, it rose from 15 million ETH to current levels, which translates to roughly $ 28 billion.

Although the idea had been circulating in other systems, Ethereum effectively reintroduced the notion of smart contracts to blockchains. Smart contracts extend the simple abstract data type notion of blockchain transactions to include complex data type classes with end-user defined variables and functions. When an end-user deploys a smart contract in a blockchain, this deployment results in instantiating an object instance of the smart contract class in the blockchain. The object state is initially stored in the block where the object is instantiated. End-users can issue a smart contract function call by sending function call requests to the miners of a blockchain. These function calls are transactions that are sent to the address of the smart contract object. Miners execute these transactions and record object state changes in their currently mined block. Therefore, the state of a smart contract object could span one or more blocks of a blockchain (Zakhary, V., Agrawal, D., & Abbadi, A. el.; 2019). Each transaction inside a block includes the sending and receiving addresses and the transferred value.

As an open shared ledger, Ethereum allows any user to store the history of the entire transaction. By using this history, special nodes (miner's node) can confirm new transactions. Miner's integrity is determined by a proof mechanism that validates miners' transactions. It notifies new transactions added to the Ethereum chain *via* blocks added at a constant rate between 10 and 20 s. Miners in the case of the Ethereum network have been essentially replaced by validators, as has been previously discussed. Ethereum is difficult to calculate when changing a transaction (double spending) that a user has already used since the processing information for all relevant blocks must be re-executed. All users of the Ethereum network receive and send transactions through ID or address generated by the Elliptic Curve Digital Signature Algorithm (ECDSA), which gives the private and public key pairs. The private key is used to send transactions to another address, and the public key is used to receive transactions from another address. Ethereum users can synchronize the nodes with the network to get information about every transaction. (Said, A., Janjua, M. U., Hassan, S. U., Muzammal, Z., Saleem, T., Thaipisutikul, T., Tuarob, S., & Nawaz, R.; 2021).

3.2.2. Smart Contracts

A smart contract is a software program integrated into a blockchain, designed to streamline contract agreements. These contracts function based on predefined conditions that all parties involved have consented to. Once these conditions are satisfied, the smart contract autonomously executes the terms of the agreement. The main form of smart contract is simple, transactional smart contracts, which normally execute the transfer of currency from one party to another. To interact with a smart contract, a user initiates a transaction to the contract's unique address on the Ethereum blockchain. This transaction may include input data, which serves as instructions for the contract. The contract processes this input according to its predefined rules. Once deployed, smart contracts are immutable. This immutability ensures that the code and rules of the contract cannot be altered, providing assurance of its integrity. Additionally, the Ethereum blockchain maintains a transparent, publicly accessible record of all contract executions and outcomes, fostering transparency and accountability.

Smart contracts are slowly becoming more detailed and advanced. The Ethereum network allows for the development of many DApps, which employ different kinds of functions and therefore varying types of coding. For example, the most common protocol, the ERC-20 token standard serves as a contractual blueprint. It outlines the essential functions and interfaces that any token wishing to adhere to this standard must implement. These functions include, but are not limited to, transferring tokens, querying token balances, and approving token allowances. Fungibility is a central tenet of the ERC-20 standard. Tokens adhering to this standard are considered interchangeable on a one-to-one basis. Each unit of a given ERC-20 token is indistinguishable from another, ensuring uniformity in their treatment across Ethereum wallets and exchanges.

The ERC-20 standard's widespread adoption has engendered compatibility among various Ethereum-based tokens. This interoperability is pivotal in enabling decentralized exchanges (DEXs), token wallets, and other blockchain applications to seamlessly support multiple ERC-20 tokens without necessitating integration for each token.

ERC-20 tokens exhibit a set of fundamental functions like "transfer", which facilitates the direct transfer of tokens between Ethereum addresses, while "balanceOf" enables the retrieval of a token holder's balance. The "allowance" function permits token holders to delegate spending permissions to others. To augment transparency and enable real-time tracking of token movements, ERC-20 tokens often incorporate event logging. This functionality logs pertinent events, such as token transfers, allowing external applications to monitor and respond to token activity on the Ethereum blockchain.

Tokens built in the ERC-20 protocol are widely embraced in Initial Coin Offerings (ICOs), token sales, and decentralized applications (DApps). Their compliance with this standardized framework simplifies the development process for token creators, fostering innovation and the creation of diverse utility tokens. It is worth noting that the Ethereum community has witnessed the evolution of token standards beyond ERC-20. These iterations, such as ERC-721 (for non-fungible tokens or NFTs) and ERC-1155 (for multi-token standards), have addressed specific use cases and advanced the tokenization landscape.

3.2.3. Ether And Tokens

Ether (ETH) is the native cryptocurrency coin of the Ethereum blockchain and the single most important crypto in terms of facilitating transactions in the whole crypto ecosystem. It is also the second largest cryptocurrency by market capitalization, standing at $ 196.577.016.892 at the time of writing, with its price at $ 1.634.35 (data by coinmarketcap.com).

Ether primarily serves as a digital currency within the Ethereum network. It is mainly used to pay for transaction fees (gas) and computational services (smart contracts) on the platform. Beyond being a medium of exchange, Ether is also a store of value and can be held as a digital asset or investment. It operates on a decentralized blockchain, meaning it's not controlled by any central authority. Instead, it relies on a distributed network of nodes for consensus and validation. Security is maintained through cryptographic principles, including public-key cryptography and proof-of-stake (although Ethereum began as a proof-of-work network). Unlike Bitcoin, Ether doesn't have a strict capped supply, however, there is a mechanism to limit its issuance rate. Since the implementation of Ethereum 2.0 and the transition to a PoS consensus mechanism, validators lock up Ether as collateral to secure the network.

The other cryptocurrencies, that operate on the network are called tokens and as has been discussed, may cover a large array of uses and representations. Some of the most important tokens, by market capitalization, are Multi-Collateral Dai (DAI), Shiba Inu (SHIB), and Polygon (MATIC). Below, we present some basic information regarding them to exemplify the variety in use cases and operations that are possible in the Ethereum ecosystem, which are functioning through and being represented by tokens.

Multi-Collateral Dai (DAI)| Market Capitalization: $ 5.345.174.477

DAI is a stablecoin operating within the Ethereum blockchain ecosystem, designed to maintain a stable value through a decentralized and algorithmic mechanism. Its unique approach to stability has garnered attention within the realm of decentralized finance (DeFi).

Utility and Functionality

DAI serves as a stable medium of exchange and store of value within the Ethereum network. Its primary utility lies in providing a stable alternative to volatile cryptocurrencies like Ether (ETH) or Bitcoin (BTC). Unlike traditional stablecoins, DAI does not rely on holding an equivalent reserve of fiat currency. Instead, it is collateralized by various crypto assets, predominantly Ether, locked in smart contracts on the MakerDAO platform.

Use Cases

The primary use case for DAI is to facilitate stable transactions and provide a reliable store of value within the Ethereum ecosystem. It is widely used in DeFi applications for lending, borrowing, and trading due to its stability. Users can generate DAI by locking up collateral in MakerDAO's smart contracts, providing a decentralized and trustless means of obtaining a stable digital currency.

Market Dynamics

DAI's value stability, typically hovering around $1 USD, is maintained through a combination of smart contracts and autonomous feedback mechanisms. The demand for DAI is influenced by the broader adoption of DeFi, as well as the health of the MakerDAO ecosystem. Market dynamics primarily focus on maintaining the peg to the US dollar while ensuring sufficient collateralization to mitigate risks.

Shiba Inu (SHIB)| Market Capitalization: $ 4.491.232.669

SHIB, often referred to as the "Dogecoin Killer," is a meme-inspired cryptocurrency that emerged in 2020. It draws inspiration from the Dogecoin (DOGE) meme culture and has gained significant attention in the crypto space.

Utility and Functionality

SHIB operates as a decentralized cryptocurrency built on the Ethereum blockchain. While its utility and functionality align with being a peer-to-peer digital currency, its primary appeal lies in its meme-driven community and the potential for speculative gains. SHIB is governed by the decentralized SHIB community, where decisions are made through token voting.

Use Cases

SHIB's use cases are primarily centred on meme culture and speculative trading. It is often used for tipping and small transactions within the SHIB community. However, its main attraction is its potential for high volatility and quick price movements, appealing to traders seeking speculative opportunities.

Market Dynamics

SHIB's market dynamics are driven by community sentiment, social media trends, and celebrity endorsements. Its value can experience rapid fluctuations influenced by the whims of traders and the broader crypto market. As a meme coin, it remains a high-risk, high-reward asset.

Polygon (MATIC)| Market Capitalization: $ 5.407.063.597

MATIC, now known as Polygon, is not exactly an Ethereum token per se but a cryptocurrency that seeks to address scalability issues on the Ethereum blockchain. It operates as a Layer 2 scaling solution, aiming to enhance Ethereum's throughput and efficiency.

Utility and Functionality

MATIC serves as the native token of the Polygon network. Its primary utility is to facilitate transactions, pay for gas fees, and participate in the network's PoS consensus mechanism. Polygon operates as a Layer 2 solution, allowing developers to build and deploy scalable DApps while benefiting from Ethereum's security.

Use Cases

Polygon's main use case is to provide a scalable platform for Ethereum-compatible DApps. Developers can leverage Polygon's infrastructure to create efficient and cost-effective DApps without compromising on security. MATIC tokens are staked to secure the network and participate in the consensus process.

Market Dynamics

The value of MATIC is closely tied to the adoption of the Polygon network. As more DApps migrate to Polygon for scalability and cost-efficiency, demand for MATIC may increase. Additionally, the network's transition from a PoS sidechain to a full-fledged PoS blockchain may impact the dynamics of MATIC's value.

Other important tokens are Uniswap, Aave, Yearn, and Compound, which broke through in the initial taking off of Ethereum as prominent DeFi platforms for lending/ borrowing and digital asset management protocols. Lastly, special mention has to be made to Tether (USDT). Tether is the number one choice of stablecoin in Centralized Exchanges (CEX). At its core, Tether serves as a digital representation of traditional fiat currencies on blockchain networks. USDT tokens are issued by a company called Tether Limited, and each USDT is supposed to be backed by a one-to-one reserve of the respective fiat currency (the US dollar), effectively providing a digital equivalent of that currency. Tether's primary use case is to provide stability in the often volatile cryptocurrency market. It serves as a safe haven where traders and investors can park their assets during periods of market turbulence. By offering a stable value that mirrors fiat currencies, USDT enables users to hedge against cryptocurrency price fluctuations. Furthermore, USDT is employed in DeFi platforms, lending and borrowing protocols and as a medium for cross-border transactions. Its value is deliberately maintained close to the value of the fiat currency it represents, typically $1 USD. The stability of USDT is achieved by backing each token with a corresponding reserve of fiat currency held in banks. Periodic audits are conducted to verify the reserve's adequacy. However, this hasn’t prevented it from facing significant controversy and legal scrutiny, primarily related to concerns about the adequacy of its reserves and allegations of market manipulation. Critics argue that the company must provide more transparent and frequent audits to verify its reserve holdings. Legal investigations have sought to establish the legitimacy of these concerns.

3.3.Decentralized Finance (DeFi)

As has been noted, Ethereum presents the possibility for the creation of many different kinds of applications. The most revolutionary aspect so far, since its inception, is probably the birth and advancement of DeFi.

The initial DeFi applications on the Ethereum network were lending and borrowing protocols like Aave (AAVE) and Compound (COMP). The way those protocols function is reminiscent of traditional forms of banking. Some users are in demand for funds and others have excess funds, upon which they would prefer to have a higher return than they would just by holding them. Normally, a third party (bank or other institution) would act as the middleman in this situation, accepting deposits of funds with the promise to pay them back plus interest and lending out funds for a higher interest rate, making a profit between the two. The notion is the same with the exception that the third party does not need to manually undertake these tasks. Instead, this role is given to a programming protocol, which automates the procedure using smart contracts. In time, these protocols have become more complicated and have the ability to execute even more advanced financial operations. Investing and asset management options have increased not only on Aave and Compound but many more DeFi DApps that have sprung up over time.

Another key aspect of DeFi is the function of DEXs. These decentralized exchanges automate the trading procedure of centralized exchanges in a similar manner to how we saw lending and borrowing being automated in the previous paragraph. However, the decentralized aspect has an even more important role in the world of crypto. Not having to rely on central authorities for the execution of economic transactions has been the goal of Satoshi Nakamoto since the inception of Bitcoin and remains so for many enthusiasts, who see DEXs as the only viable way of transacting in the crypto economy. Notwithstanding the aspect of automation, CEXs retain the wallets' private keys in their possession, unlike what occurs when transacting on DEXs and as the crypto adage says "not your keys, not your coins".

Special mention has to be given to another integral part of DeFi, which unlike the previous examples is not just a decentralization of common practices in traditional finance but more unique in the cryptocurrency ecosystem. We're referring to automated market makers (AMMs), which have helped immensely in the widespread adoption on the use of DEXs. Before AMMs, limited liquidity posed a challenge on DEXs, hindering regular trading. AMMs address this issue by forming liquidity pools (LPs) and incentivizing liquidity providers to contribute assets to these pools. A smart contract with a mathematical formula is used to price assets, and users trade against the liquidity locked within smart contracts rather than relying on traditional order books. Liquidity providers can earn transaction fees that the protocol levies for carrying out transactions through liquidity mining (Shah, K., Lathiya, D., Lukhi, N., Parmar, K., & Sanghvi, H.; 2023). Each LP is comprised of a pair of cryptocurrencies that providers deposit into the pool. In exchange, they receive a special token, called a LP token, which serves as receipt to an account. When providers wish to withdraw their capital, they re-exchange the LP token for the original cryptocurrencies with the token getting “burnt” in the process. Users interact with the pool by exchanging one currency for another while also paying a small transaction fee. When someone wants to trade, pairs behave as AMMs, ready to take one token in exchange for the other as long as the “constant product” formula is followed. Assume that Alice purchases 1 ETH for 300 USDT from the ETH/USDT liquidity pool. By doing so, she raises the USDT share of the pool while decreasing the ETH portion. This essentially indicates that the price of ETH rises. After the transaction, the pool has less ETH, and the overall liquidity (k) must remain constant (Shah, K., Lathiya, D., Lukhi, N., Parmar, K., & Sanghvi, H.; 2023). Uniswap (UNI) is the number one LP protocol in Ethereum and the example above fully reflects the way it operates.

3.4. On-Chain Metrics

The function of the blockchain with its app building, interacting, transacting and all other activities that occur in or through it, generates some important and quantifiable pieces of information, in the form of what we call on – chain metrics. Such information can be harnessed and used in several fields of data analysis and machine learning to provide a vast array of insights. The following table provides a list of some basic metrics, which are commonly searched by different kinds of interested parties.

Table 2: Basic Ethereum on-chain metrics

Transaction Count	The total number of transactions processed on the Ethereum network within a specific time frame.
Block Height	The current block number in the Ethereum blockchain, indicating the chronological order of blocks.

Gas Used	The total amount of gas consumed by all transactions and smart contracts within a block.
Uncle Rate	The rate at which uncle (or stale) blocks are generated, which indicates network congestion and mining competition.
Difficulty	A measure of how hard it is to mine a new block, which adjusts periodically to maintain block generation time.
Hashrate	The total computational power (measured in hashes per second) dedicated to Ethereum mining.
Block Time	The average time it takes to mine a new block, which is designed to be around 15 seconds in Ethereum.
Gas Price	The average price (in Ether or Gwei) users are willing to pay per unit of gas for transactions, indicating network congestion.
Active Addresses	The number of unique addresses that have conducted transactions or interactions with smart contracts on the Ethereum network.
Token Transfers	The total number of transfers involving Ethereum-based tokens, providing insights into token activity.
Token Holders	The number of unique addresses holding specific tokens, offering a view of token distribution.
Total Supply	The total supply of Ether (ETH) or specific tokens in circulation.
Gas Limit	The maximum amount of gas allowed in a block, which affects the capacity of the Ethereum network.
Gas Used Percentage	The percentage of the gas limit utilized in a block, indicating network congestion levels.
Miner Revenue	The total rewards earned by miners, including block rewards and transaction fees.
Smart Contract Deployments	The number of new smart contracts deployed on the Ethereum blockchain.

Token Balances	The amount of specific tokens held in Ethereum addresses, providing insights into token distribution.
Token Transfers Per Day	The daily count of token transfers, offering a dynamic view of token activity.
Active DApps	The number of active decentralized applications (DApps) running on the Ethereum network.
DeFi Metrics	Metrics related to decentralized finance (DeFi) protocols, including total value locked (TVL), loan volume, and liquidity pools.
Average Transaction Value	The average value of Ether (ETH) transferred in each transaction on the Ethereum network, providing insights into the size of transactions.
Token Velocity	A measure of how quickly a token is changing hands within the network, calculated as the total transaction volume divided by the token's market capitalization.
Token Age Consumed	The total number of days since the tokens in a transaction were last moved, helping to identify long-term holders and short-term speculators.
DApp Usage	Metrics related to decentralized application (DApp) usage, including the number of users, transactions, and volume on popular DApps.
DEX Metrics	Metrics specific to decentralized exchanges, such as trading volume, liquidity, and the number of trading pairs.
Token Prices	The price of Ethereum-based tokens, including historical price data, market capitalization, and trading volume.
Gas Efficiency	Measures how efficiently gas is used in smart contracts and transactions, helping developers optimize their code.
Active Validators	The number of active validators participating in Ethereum's proof-of-stake (PoS) network.
Block Size	The size of Ethereum blocks in bytes, which affects the network's scalability and capacity.

Token Age Distribution	A breakdown of token holders based on the age of their holdings, showing the distribution of long-term and short-term holders.
Contract Calls	The number of interactions with smart contracts, including calls, deployments, and interactions with DApp protocols.
DAO Metrics	Metrics related to DAO (Decentralized Autonomous Organization) activity, including voting participation, proposals, and token holdings.
NFT Metrics	Metrics specific to NFTs, such as the number of unique NFTs, trading volume, and top NFT collections.
Gas Price Trends	Historical data on gas prices, showing trends in gas costs over time.
Staking Metrics	Metrics related to staking in Ethereum 2.0, including the total amount of ETH staked and staking rewards.
Transaction Confirmation Time	The average time it takes for a transaction to be confirmed on the Ethereum network, which can vary based on network congestion.

Metrics like these are used for a variety of analytical purposes. As we are more interested in financial types of analysis, we provide some elementary examples of their potential usage. E.g.

1. **Token Age Consumed:**

   Token Age Consumed measures the movement of previously dormant tokens. A significant increase in Token Age Consumed may indicate that long-term holders are selling or moving their tokens, potentially signalling a shift in sentiment and affecting price trends.

2. **Active Addresses:**

   Monitoring the number of active addresses can provide insights into user engagement and network activity. A surge in active addresses may correlate with increased demand and potentially influence price movements.

3. **Token Transfers:**

   An increase in the number of token transfers may suggest higher token usage for transactions or interactions with decentralized applications. This increased utility can impact the token's value and price trends.

4. **Gas Usage:**

   Gas usage can reflect the overall activity and demand on the Ethereum network. A significant increase in gas usage may indicate heightened network congestion, potentially affecting transaction times and costs, which can, in turn, influence token prices.

5. **Token Velocity:**

   Token Velocity measures the rate at which tokens are changing hands within the network. A high token velocity may suggest increased trading activity and shorter holding periods, potentially affecting price volatility.

6. **DEX Trading Volume:**

   Decentralized exchanges (DEXs) often provide on-chain trading volume data. An increase in DEX trading volume may signal heightened trading activity, affecting token prices, especially for tokens listed on these exchanges.

Of course, there are even more metrics to take into account, like **New Addresses** (on a particular day or other specified time period), **Zero Balances All-time** (sum of zero-balance addresses since inception), **Unique Addresses All-time** (sum of addresses with at least one transaction since inception) , **Hash Rate** (the hash rate for a day is the average difficulty / the average time between blocks for the day / 10^{12} and it is expressed in TH/s [1,000,000,000,000 hashes per second]) or **Current Supply** (current supply refers to the average size in bytes of all blocks created that day). Even more particularized and advanced metrics can be found and exploited, such as **Burned ETH after EIP-1559** or **Net ETH Emission after EIP-1559** depending on the specific goal of each individual, who aims to analyse them.

A large part of blockchain originated information, that's considered of severe importance to analysts and investors revolves around capital movements. There are metrics that follow UTXO data from transactions, which calculates the sum of unspent funds after a transaction. This data originates from a process called price-stamping (recording the market price or value of a cryptocurrency at the time a UTXO is created or when it becomes unspent. This information is typically stored alongside the UTXO data). This way, the length of time funds remain in wallets can become identifiable, thus generating an image about the current situation of HOLDing strategies and investor sentiment. While this was a popular idea for tracking coin information in Bitcoin, it is not available within Ethereum's network, as it uses an account-based system. Many of these metrics have been mapped on to Ethereum's system

system, like the **SOPR** (Spent Output Profit Ratio), which is calculated by dividing the realized value of a spent output by the value at creation of the original UTXO, by (Kilian Heeg, & Rafael Schultze-Kraft; 2020) and others.

One of the most prominent metrics in the field is **Realized Capitalization**, a metric that has become increasingly popular in recent years and has given birth to even more metrics that use it as its basis. It is calculated by multiplying the price at which each unit of a cryptocurrency (e.g. Ether) was last transacted (last traded price) with the number of units that have been moved since that last transaction. In essence, Realized Capitalization takes into account the market value of coins based on when they were last actively traded or moved on the blockchain, rather than the current market price. This metric is useful for assessing the distribution of a cryptocurrency's value among its holders and understanding the price levels at which a significant portion of the coins changed hands. Realized capitalization can be particularly valuable in analysing long-term holders' behaviour and identifying price levels where investors initially acquired their coins. It provides insights into the potential support and resistance levels based on historical price action, which can be different from current market prices.

Metrics originating from or comprising elements of Realized Capitalization encompass **Unrealized Profit and Loss** metrics, as well as **Realized Profit and Loss** metrics. These analytical instruments are especially well-suited for conducting market analyses and formulating trade strategies, as they effectively signal the underlying financial incentives, namely fear and greed that guide investor decision-making. For example, as is mentioned in Glassnode's analytical article "The Foundational On-chain Metric: The Realized Cap" (Checkmate; 2023), Realized Profits tend to dominate during macro uptrends and frequently peak around local and global market highs. This describes the mechanic of profit taking which eventually over-saturates incoming demand. Realized Losses on the other hand are known to prevail during macro bearish trends and have historically experienced two notable capitulation spikes, as they are frequently called. One at the beginning of a bear market and one as the market approaches maximum negative sentiment and eventual exhaustion.

An important metric, which actually combines the above in its calculation, is **NUPL** (Net Unrealized Profit and Loss). This metric measures the difference between the realized value (used interchangeably with realized capitalization) and the current market value (the current price at which the currency is trading). When NUPL is high (greater than 1), it suggests that a significant portion of holders is in a state of profit, potentially indicating bullish sentiment and an increased likelihood of profit-taking or selling pressure. Conversely, when NUPL is low (less than 1), it indicates that many holders are at a loss, suggesting bearish sentiment, but it may also indicate potential buying opportunities.

The final mention of Realized Cap-based on-chain metrics will be given to **Market Value to Realized Value Ratio** or **MVRV**. The MVRV Ratio was first proposed by (Mahmudov, M., & Puell, D.; 2018) in an article on medium.com and is easy to calculate since it’s a simple division of the market value of an asset by its realized value on a daily basis. When the ratio is more than 1, it implies overvaluation of the cryptocurrency in question while values below 1 imply undervaluation. The aimed use of the MVRV Ratio is mainly to detect accumulation and distribution phases, which in the crypto ecosystem effectively represent a coin’s market cycle.

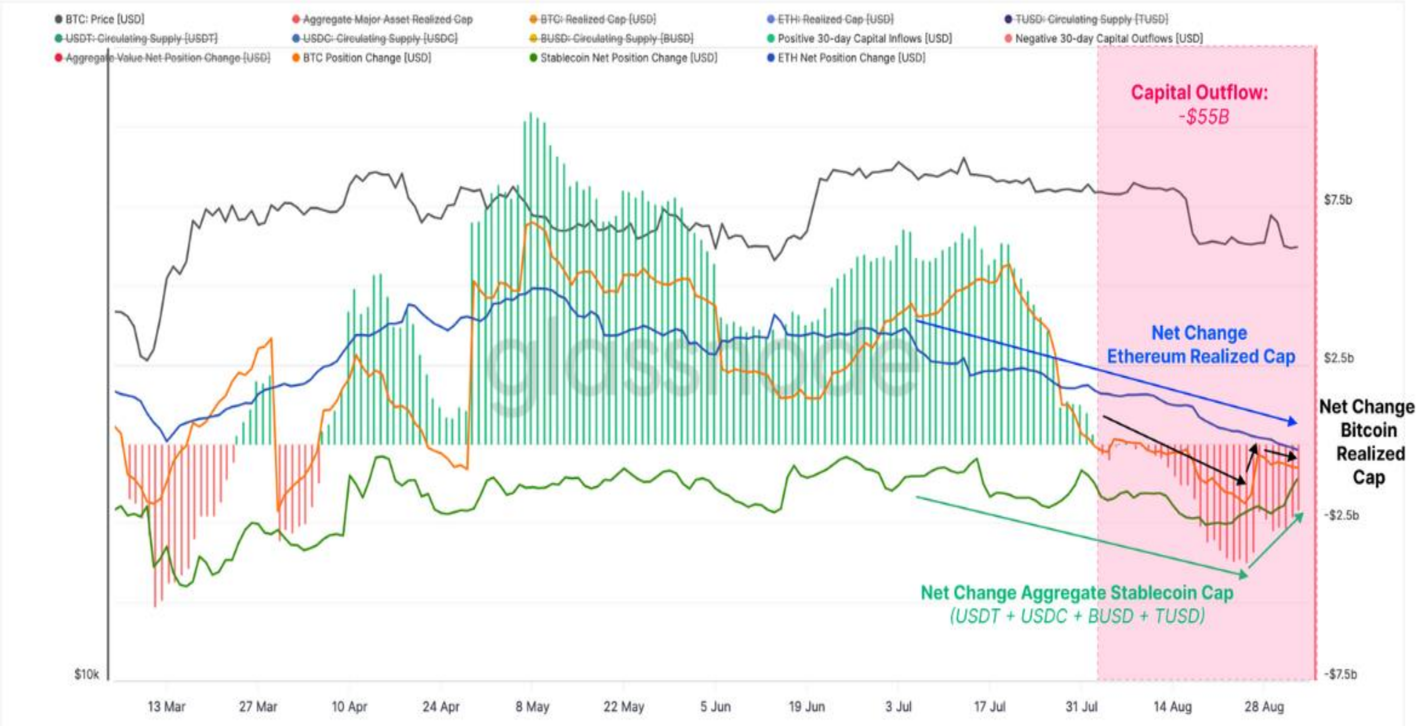
4. METHODS OF ANALYSIS

4.1.Market Applications of On-Chain Analysis

On-chain metrics are commonly used in analysis of crypto markets from professionals or even non-professional traders and investors. They are increasingly becoming a popular tool to extract useful insight, with the intention of using that to guide investor actions. In the previous section we defined and discussed the origin of these metrics. We also provided examples of such data and saw a handful of ways, in which they can be used by interested parties. In this section, we will examine real examples of analysis using on-chain data, which is used daily by investors and others to guide their decision making.

4.1.1. Aggregate Realized Value And Capital Flows

The first example comes from the weekly edition of Glassnode's newsletter (Glassnode is one of the largest online platforms for crypto analytics, especially regarding on-chain metrics and relevant data). The analyst (Kohn, A.; 2023) makes use of the Realized Capitalization metric that we've mentioned or rather, one of its by-products. The **Aggregate Realized Value** metric combines the Realized Cap of BTC and ETH with the supply of the five leading stablecoins, namely USDT, USDC, BUSD, DAI, and TUSD. In its essence, it's a simple addition of the aggregated realized cap of the two main coins and the aggregated supply of the stablecoins. {ARC(BTC,ETH) + AS(stablecoins) = ARV(market)} An increase in either part of the equation would indicate an increase in capital inflows in the market, the combination of which provides an accurate depiction of net inflows/outflows in the general crypto market. Of course, the market is comprised of many more altcoins and tokens but the logic behind this approach is that investors either buy BTC and ETH directly or allocate funds into altcoins through acquiring one of the five most popular stablecoins first. In this example, the analyst concluded that the market entered a state of capital flight in the month of August 2023, with approximately $ 55 billion leaving the crypto space.



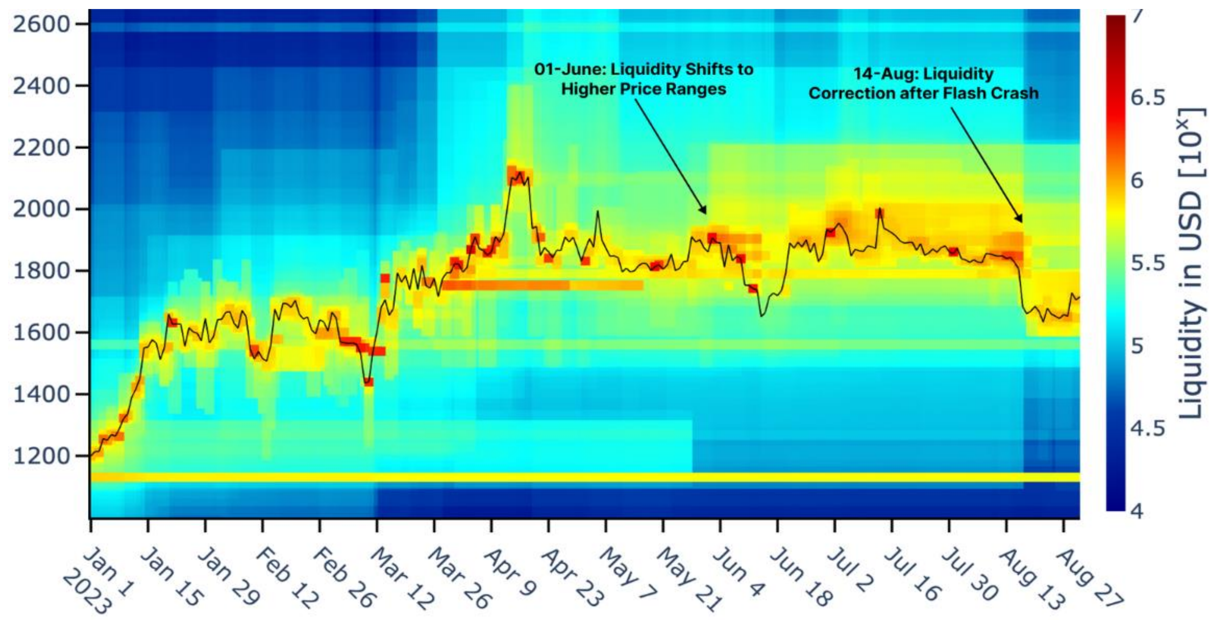
Graph 1: Aggregate Market Realized Value net position change (source: glassnode.com)

4.1.2. Market Sentiment Analysis Through Liquidity Pool Data

On the same publication, another interesting example appears. In this instance, the author's intention is to track investor sentiment through liquidity pools in Uniswap V3. Before we go deeper in the analysis, we have to mention that she seeks to identify pricing information from LPs, in a similar manner to options markets. This is based on a hypothesis proposed by (Lambert, G.; 2021) in the medium.com article, titled "Uniswap V3 LP Tokens as Perpetual Put and Call Options". Without ascribing fault to this particular approach, it must be noted that it is not part of any peer reviewed research and therefore we cannot fully condone its conclusions. However, (Deng, J., Zong, H., & Wang, Y.; 2023) have also noted the similarities between providing liquidity in LPs and options trading. Specifically, due to the concept of concentrated liquidity provision (the ability for providers to concentrate their assets within a specific price range in a LP, optimizing capital efficiency), that was introduced in the V3 version of Uniswap, they characterize the structure of impermanent loss in the protocol as "option-like".

In this particular analysis, the author chooses to focus on the USDC/ETH 0.05% LP on Uniswap, as it is considered the most active in the protocol with a 7-day trade volume of $ 1,51 billion and a TVL of $ 260 million. The metrics used are **mints per day** and **burns per day**, referring to the minting of new LP tokens and the burning of previous ones respectively. The first one represents the number of liquidity positions opened and the latter the number of closed positions. The net change in LP positions is regarded as less affected by market trends than by short-term volatility, like that caused by particular events. The examination of the data shows that approximately 30.4% of the total liquidity is located within an 11% price range, with expected downside of -2.7% and upside of +8.6%. At the same time, another large portion of the liquidity is positioned within a -8.5% to +23.7% buffer of the current ETH price ($ 1.642). Interestingly, this distribution of positions largely aligns with option strike prices for contracts expiring at the end of September. In particular, the percentage of calls that have a strike price between $ 1.700 and $ 2.300 is estimated at 70% while 75% of puts have a strike between $ 1.300 and $ 1.900. Due to the adoption of advanced automated liquidity LP strategies and execution methods, liquidity providers have been reported to achieve notable success in strategically positioning liquidity very close to the prevailing market price, especially during periods of heightened volatility. Notably, on 1/6/23, a substantial volume of liquidity was strategically placed just above the prevailing price level at that time, as indicated by the deeper yellow region of the heatmap presented below (the density of liquidity is expressed in increasingly cool-to-hot colours). This could reasonably be interpreted as an anticipation of higher fee earnings by market makers in this particular price range. This liquidity configuration remained in effect until the occurrence of a flash crash in August, at which point liquidity concentrations were adjusted to be progressively situated below the $ 1.800 mark. Additionally, the analyst makes the observation that significant concentrations of liquidity, represented by the red regions, tend to align with strong price movements and frequently coincide with reversals in market sentiment.

Effectively, the analysis concludes that LP position data from pools, which employ the concentrated liquidity provision, can provide useful insight regarding market sentiment, similarly to the observation of options contracts positions.



Graph 2: Liquidity concentration in LPs heatmap (source: glassnode.com)

4.2. On-Chain Analysis In Academia

With the increasing interest taken by economists and other academic scholars and scientists in crypto, we are starting to see analytical and empirical approaches on crypto assets that employ on-chain data. The subjects of price prediction and asset valuation have always been flagships in the field of financial research, therefore in the following paragraphs, we take a closer look in some academic examples of research using on-chain data and analytical methods in order to acquire a better understanding of these new and dynamic approaches on an otherwise classic field of finance.

4.2.1. Price Prediction Of ETH And Ethereum Tokens

We begin with a few empirical models that attempt to capture the value of cryptocurrencies, specifically ETH along with some other Ethereum based tokens. The first approach we're going to examine is a study by (Bakhtiar, T., Xiaojun, L., & Adelopo, I.; 2023). The research is not exactly particular to Ethereum, as not all of the 42 selected cryptocurrencies are part of the Ethereum network and it's not based on a completely on-chain approach. However, we have chosen to include it as it serves as a kind of bridge on the shift from more traditional methods to on-chain analysis. Its aim is to estimate the impact that fundamental factors and general sentiment have on the valuation of cryptocurrencies. We are mostly interested in what the researchers considered as fundamental factors, considering the nature of cryptocurrencies, so we will focus on the testing of the first hypothesis in the research (the other two include regressions against the Fear and Greed Index and Google Search Index).

For this purpose, they conduct a panel regression model statistical analysis, in which daily return on investment (RTN) of each currency is considered as the dependent variable. The independent variables are then comprised of the adoption of a PoW consensus mechanism, the adoption of a PoS consensus mechanism, the presence of smart contracts (SC), whether there has been an initial coin offering or not (ICO), the adoption of an inflationary policy with unlimited supply (IS) and the adoption of limited supply policies (LS). It is plain to see that the variables that the researchers deemed as fundamental factors for price determination in cryptocurrencies mainly stem from the form and function of the blockchain, which each cryptocurrency is a part of. Additionally, tokenomics decisions such as those in the last two variables also seem to be of fundamental importance. The only fundamental factor, which isn't determined from facts regarding the blockchain, is the choice of an ICO or not for the release of the token in question. Finally, they add a control variable for whether a token is considered a stablecoin or not (NSC). The form of the regression formula becomes like this:

$$\mathrm{RTNit}=\beta 0+\beta 1\mathrm{PoWit}+\beta 2\mathrm{PoSit}+\beta 3\mathrm{SCit}+\beta 4\mathrm{ICOit}+\beta 5\mathrm{ISit}+\beta 6\mathrm{LSit}+\beta 7\mathrm{NSCit}+\varepsilon\mathrm{it},$$

where $\beta 0$ is the constant and εit is the error term.

The findings of the regression indicate that the choice of the consensus mechanism as well as the presence of an ICO are the main driving forces behind daily returns while the impact of the other factors is not statistically significant for non-stablecoins. For the record, the results of the statistical model are presented in the table below.

Table 3: Cryptocurrency fundamental factor regression results					
	Independent variables	Coefficient	Std error	t-statistic	p-value
Training of non-stable coins	PoS	0.0019	0.001	2.087	0.037**
	PoW	0.0022	0.001	3.083	0.002***
	Smart contract	-0.0004	0.001	-0.389	0.697
	ICO	0.0026	0.001	2.684	0.007***
	Inflationary supply	0.0009	0.001	1.113	0.266
Testing of non-stable coins	PoS	2.657×10^{-5}	0.001	0.019	0.985
	PoW	0.0036	0.001	3.260	0.001***
	Smart contract	-0.0010	0.001	-0.656	0.512
	ICO	0.0044	0.002	2.581	0.004***
	Inflationary supply	0.0005	0.001	0.448	0.654
Training of stable coins	PoS	0.0016	0.001	1.239	0.215
	PoW	0.0006	0.002	0.351	0.726
	Smart contract	0.0004	0.002	0.152	0.879
	ICO	-0.0003	0.001	0.241	0.810
	Inflationary supply	-0.0009	0.001	-0.681	0.496
Testing of stable coins	PoS	-0.0015	0.002	-0.853	0.394
	PoW	-0.0008	0.003	-0.317	0.751
	Smart contract	0.0012	0.003	0.345	0.730
	ICO	-0.0017	0.002	-0.901	0.368
	Inflationary supply	-0.0034	0.002	-1.943	0.052**

where *, **, *** represent significance at 10%, 5%, and 1%, respectively.

The next instance is focused on on-chain data in particular, which is also imported into econometric models. In their dissertation, (Stober, A., & Sandner, P.; 2020) noticed the lack of non-bitcoin related crypto studies thus far in the literature as well as the non-use of transactional data from the blockchains. They therefore set out to analyse a variety of ERC-20 tokens and determine the leading factors of price formation.

They chose to utilize an Autoregressive Distributed Lag model because of their assumption that there might be non-stationarity in the time series:

$$y_t = c_0 + c_1 t + \sum_{i=1}^{p} \emptyset_i y_{t-i} + \sum_{i=0}^{q} \beta_\iota x_{t-i} + \epsilon_t.$$

The model is applicable in the case that the non-stationary variables are not cointegrated, which means there is no long term equilibrium but only short term relationships. If the stationary variables (in level form and first difference) exhibit cointegration, it is understood that there is a long term relationship between the variables that's interrupted by short term dynamics in play, which cause deviations from the equilibrium. In this case, the ARDL is of no use and an Error Correction Model is applied instead:

$$\Delta y_t = c_0 + c_1 t + \alpha(y_{t-1} - \theta x_t) + \sum_{i=0}^{p-1} \psi_{yi} \Delta y_{t-i} + \sum_{i=0}^{q-1} \psi'_{xi} \Delta x_{t-i} + \epsilon_t.$$

This approach was judged as preferable due to a number of factors, such as that endogeneity becomes subordinated when all variables are assumed to be endogenous. Also, the model is able to differentiate between dependent and independent variables in the presence of a single long-term relationship while identifying cointegrating vectors and finally, the ECM can be derived from the ARDL.

The independent variables the researchers decided to examine are the monthly active addresses (expressed logarithmically), the natural logarithm of daily BTC closing prices (which was chosen as a dummy for the general direction of the market due its dominance in the crypto space and its common pairing with other assets in trading), as well as the ratio of tokens sent to exchanges by the sum of total tokens exchanged on that day, the ratio of tokens received from exchanges by the same denominator, and the monthly active addresses of stablecoins (expressed logarithmically). Of course, the dependent variable is none other than the token's market price (natural algorithm).

After running bounds and cointegration tests, they found that six of a total of nineteen tokens exhibited cointegration, namely Chainlink (LINK), Basic Attention Token (BAT), Ox (ZRX), KuCoin Shares (KCS), Aeternity (AE), Status Network (SNT). They then proceeded with the estimation of the six, yielding the results presented in the table in the following page.

Ultimately, the results indicate that the number of monthly active addresses of the analysed tokens demonstrate a long-term connection with their price. In five cases, the monthly active addresses of stablecoins exhibit a negative long-term relationship with the analysed tokens' prices. The selected metrics related to tokens being sent to and received from exchanges generally do not exert a significant impact on the market price of the target token while on the other hand, BTC, when considered as a market proxy, shows a positive association with all the examined tokens, except for LINK, over the long term.

Table 4: ERC-20 tokens estimation

ERC-20 token	LINK	BAT	ZRX	KCS	AE	SNT
	Log (market price)					
Long-term estimation						
Log (monthly active addresses)	0.634***	0.365***	0.352***	-0.054	0.520**	0.746***
	(0.080)	(0.094)	(0.061)	(0.073)	(0.219)	(0.103)
Tokens-to-exchange	-1.457**	0.512	-0.117	-0.841**	1.314	0.610**
	(0.557)	(0.706)	(0.164)	(0.388)	(0.834)	(0.297)
Tokens-from-exchange	0.906**	1.051*	0.595**	-0.064	-0.601	0.724
	(0.436)	(0.585)	(0.276)	(0.312)	(0.587)	(0.505)
Log (BTC market price)	0.252	0.526***	0.634***	1.531***	0.994***	0.560***
	(0.161)	(0.169)	(0.050)	(0.271)	(0.280)	(0.128)
Log (monthly active addresses of stablecoins)	0.017	-0.191***	-0.367***	-0.199***	-0.346***	-0.274***
	(0.047)	(0.055)	(0.017)	(0.070)	(0.095)	(0.034)
Short-term adjustment						
Δ Log (market price)$_t$	-0.119***		-0.085*	-0.183***		
	(0.040)		(0.044)	(0.040)		
Δ Log (monthly active addresses)$_t$	1.036***	0.360***	0.169***	0.040***	0.182**	0.339***
	(0.117)	(0.130)	(0.054)	(0.012)	(0.078)	(0.085)
Δ Log (monthly active addresses)$_{t-1}$		0.367***		-0.025*	0.170**	
		(0.140)		(0.013)	(0.079)	
Δ Tokens-to-exchange$_t$	0.119***				-0.040*	
	(0.034)				(0.023)	
Δ Tokens-to-exchange$_{t-1}$	0.067**				-0.045**	
	(0.028)				(0.021)	
Δ Tokens-from-exchange$_t$	-0.092***		-0.073**			-0.050**
	(0.028)		(0.028)			(0.024)
Δ Tokens-from-exchange$_{t-1}$	-0.075***		-0.073***			
	(0.023)		(0.024)			
Δ Log (BTC market price)$_t$	0.972***	1.003***	0.874***	0.988***	1.087***	0.993***
	(0.068)	(0.055)	(0.062)	(0.069)	(0.052)	(0.053)
Δ Log (BTC market price)$_{t-1}$	0.129*	-0.125**		-0.200**		
	(0.078)	(0.056)		(0.079)		
Δ Log (stablecoin monthly active addresses)$_t$			-0.347**			
			(0.159)			
Δ Log (stablecoin monthly active addresses)$_{t-1}$		-0.207***				
		(0.073)				
Constant	-0.690***	-0.414***	-0.779***	-0.344***	-0.296*	-0.786***
	(0.224)	(0.140)	(0.235)	(0.115)	(0.172)	(0.165)
Diagnostics						
N	571	571	315	571	571	571
F value	20.46***	22.74***	21.00***	21.55***	35.99***	25.42***
Adjustment factor	-0.081	-0.049	-0.127	-0.033	-0.029	-0.062
R2	0.400	0.440	0.530	0.384	0.493	0.467
Adjusted R2	0.381	0.420	0.505	0.366	0.479	0.449
Durbin-Watson test	2.012	2.092	1.941	1.980	1.919	1.983

*Significance levels: * 0.05 < p ≤ 0.10; ** 0.01 < p ≤ 0.05; *** p ≤ 0.01. For the short-term adjustment: Empty cells are not significantly different from zero.*

4.2.2. Asset Valuation Using Neural Networks

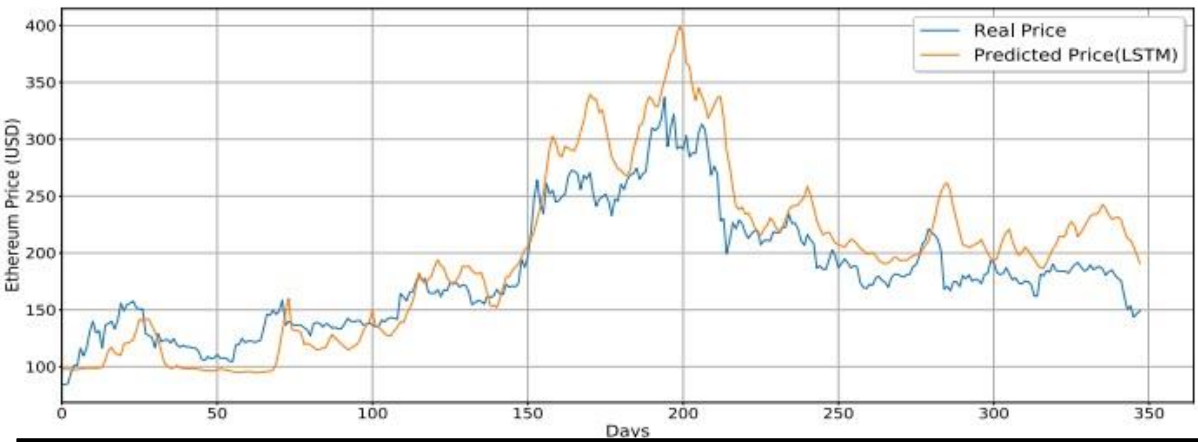
In the previous section, we discussed some examples of relatively simple econometric analysis of on-chain data using different types of regressions, in order to predict the price of ETH, along with some other coins, and various ERC-20 tokens. In this section, we are going a little deeper and we are investigating approaches, which incorporate neural network models along with on-chain metrics, for the same purpose of price prediction.

Blockchain data is timestamped, embedded in open ledgers and immutable. This creates the possibility for advanced predictive models to identify trends with greater ease and precision, compared to other data sets. Such is the opinion of (Jagannath, N., Barbulescu, T., Sallam, K. M., Elgendi, I., McGrath, B., Jamalipour, A., Abdel-Basset, M., & Munasinghe, K.; 2021). The above undertake the effort to combine more advanced algorithms and deep learning techniques with on-chain data to increase accuracy in price prediction models, such as the Long Short Term Memory Recurrent Neural Network or LSTM-RNN, which is a specific type of recurrent neural network capable of solving both long-term and short-term dependence issues in a resilient and efficient manner. In order to select which metrics to employ, they ran correlation tests (Pearson's and Spearman's methods) between different on-chain metrics and normalized ETH prices and picked the ones that seemed more strongly associated with price movements. They inserted a variety of data, such as the size of the blockchain, the number of blocks linked to it, or the difficulty of mining blocks from the Ethereum public blockchain and API of online resources in a normalized form. The totality of the data is presented on the following table 5.

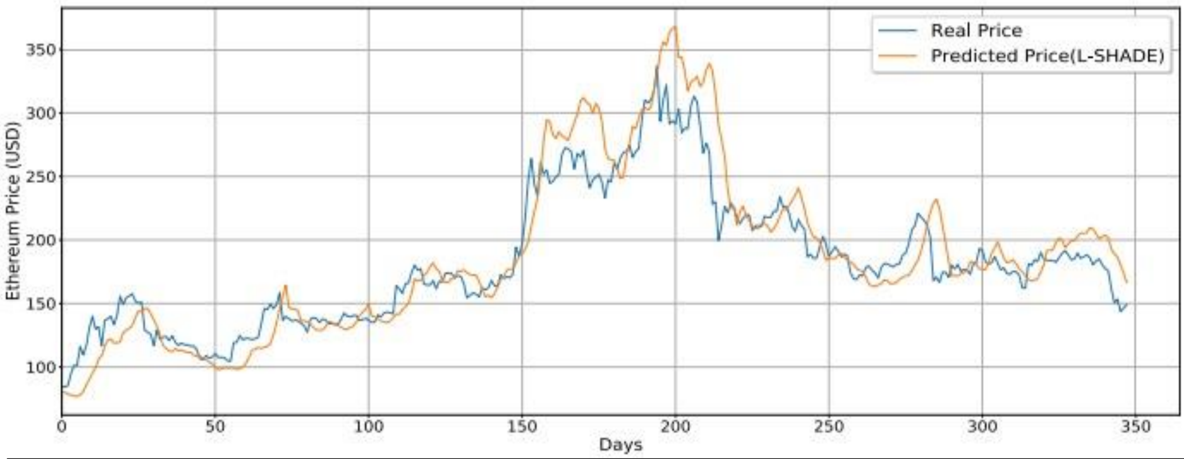
They also used three unique self-adaptive techniques (L-SHADE, jSO, MPEDE) to adjust the seventeen hyperparameters involved in the RNN architecture. We will not be going into details about the construction of these algorithms as they are outside the scope of our study. The proposed algorithms produced 150 distinct models, each with its own set of parameters and eventually, the optimal parameter combination with the most negligible loss was selected to develop an Ethereum prediction model. The training size for the algorithms was 60% of the sample, testing on the remaining 40%. Below are presented the results of the study, in which the LSTM models with the self-adaptive algorithms using on-chain metrics can be compared among each other and against a traditional LSTM model. After conducting mean absolute error (MAE), mean square error (MSE) and mean absolute percentage error (MAPE) comparison between the different self-adaptive techniques and the traditional LSTM model, it is unveiled that all three variations outperform the traditional model, with the lowest error rate belonging to the L-SHADE-LSTM model.

Table 5: On-chain metrics correlation to ETH price

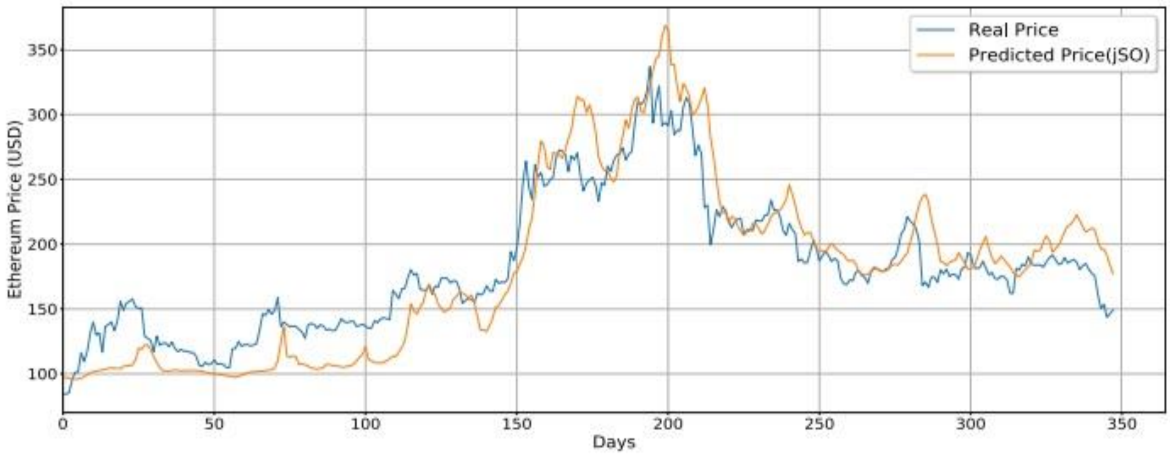
Data Category	Pearson Correlation	Spearman Correlation
Block Size	0.7465	0.5494
Block Height	0.4483	0.1640
Hash rate	0.8791	0.5681
Difficulty	0.8865	0.5663
Transactions Volume	0.3892	0.4213
Transactions rate	0.7651	0.7723
Gas price	0.4420	0.6019
Total Gas used	0.6038	0.3875
Supply in Smart Contracts	0.8725	0.6153
Internal Contract Calls	0.6218	0.3549
External Contract Calls	0.5128	0.3549
Miners Revenue	0.5924	0.4440
Miners Inflow	0.1335	0.3191
Miners Outflow	0.0612	0.2136
Miners to exchanges	0.0376	0.1904
Exchange Withdrawls	0.3943	0.5902
Exchange Deposits	0.1778	0.3161
Exchange Inflow	0.1589	0.3427
Exchange Outflow	0.0156	0.1139
Exchange NetFlow	0.4292	0.2105
Daily Active Address	0.7889	0.7277
Total Address	0.5297	0.1640
Total New Address	0.6509	0.6856
Addresses (> 1 coin)	0.5872	0.5365
Addresses (> 10 coins)	0.6215	0.5912
Addresses (> 100 coins)	0.4545	0.4163



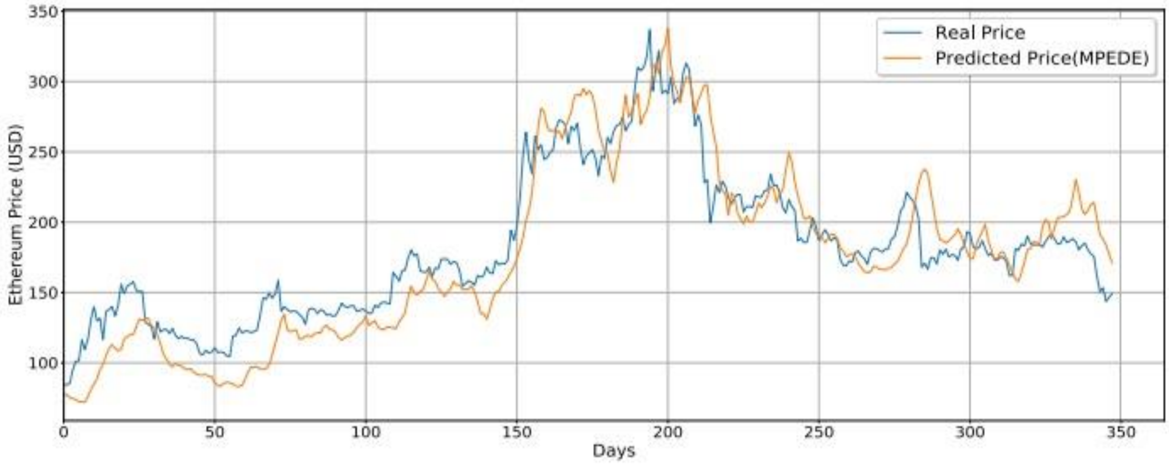
Graph 3: Real vs predicted price using a traditional LSTM model



Graph 4: Real vs predicted price using L-SHADE-LSTM model



Graph 5: Real vs predicted price using jSO-LSTM model



Graph 6: Real vs predicted price using MPEDE-LSTM model

4.2.3. Impermanent Loss Analysis On Uniswap V3 LPs

This segment of our study focuses on a subject that we briefly touched upon in section 4.1.2, where we discussed the use of data from V3 LPs on Uniswap to make judgements about the market sentiment in similar manner, to which researchers would apply to options positions in traditional finance. In the following paragraphs, we are going to look into two analyses regarding the phenomenon of impermanent loss (IL). Impermanent loss is a temporary decrease in the value of assets that liquidity providers may encounter when they supply assets to a LP in a DEX or other AMM. It's called "impermanent" because this loss only occurs on paper and isn't realized until the assets are withdrawn from the pool. The loss typically occurs because AMMs rebalance asset holdings in response to price changes caused by trading activity. If prices eventually return to their original levels, the impermanent loss diminishes or disappears. The first research analyses the phenomenon itself and tries to make sense of the operation of the LP system while also attempting to estimate the profitability of LPs in light of IL. The second one's aim is to provide possible hedging strategies to the risk of IL in open LP positions, using options contracts. Although the first of the two examples is not part of a scientific journal, rather part of a financial consultants' firm periodical, we believe it matches perfectly with the second one, hence we decided to present them in a common section.

(Loesch, S., Hindman, N., Welch, N., & Richardson, M. B.; 2021) begin their analysis by defining AMMs and IL as we have already done. They point out the fact that IL hurts liquidity providers not only in the event of one of the currencies held losing its value but also in the event of the value rising because in this case, the liquidity provider would have a greater return if they simply held the asset in their wallet. The mathematical definition of an AMM's curve is then given as the following with k as the pool constant, and m and n the constituents, each expressed in their own numenaire:

$$\mathrm{k} = \mathrm{m} \cdot \mathrm{n}$$

The authors attempt to create an accurate mathematical expression of IL by defining $\mathrm{x_t}$ as the change in the price ratio of ETH – USDX in a hypothetical pool and HODL as a portfolio with a simple holding strategy. Its value is then given as:

$$\mathrm{HODL(x)} = 1 + \mathrm{x}$$

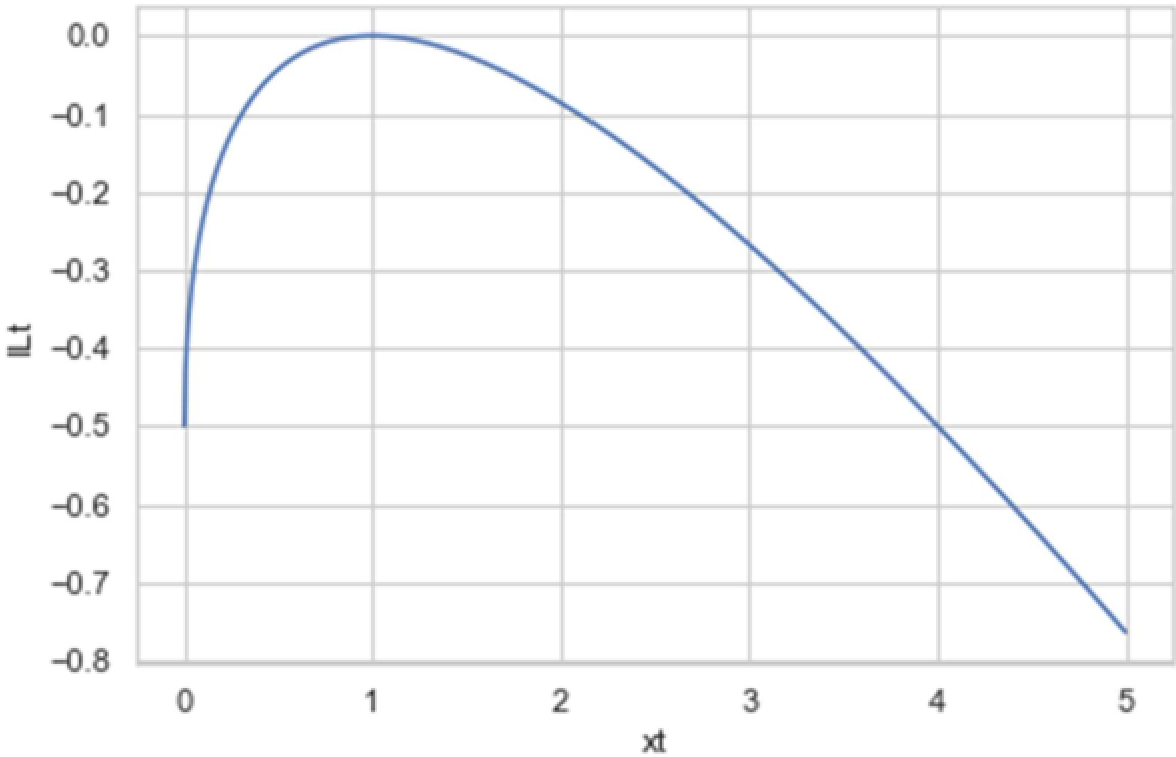
They also consider as a given, the equation for the value of the AMM portfolio:

$$\mathrm{AMM(x)} = 2\sqrt{\mathrm{x}}$$

Eventually, the IL equation is produced:

$$IL_t = \frac{AMM_t - HODL_t}{HODL_0} = \frac{1}{2}(2\sqrt{x_t} - x_t - 1)$$

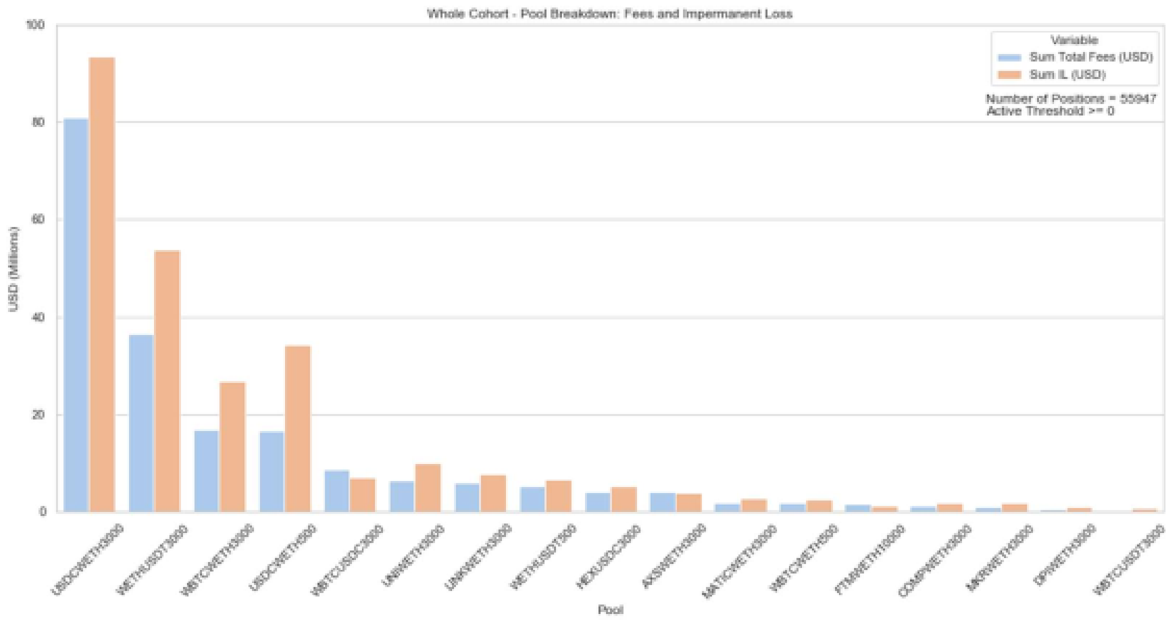
The following curve represents diagrammatically the relationship between IL_t and x_t.



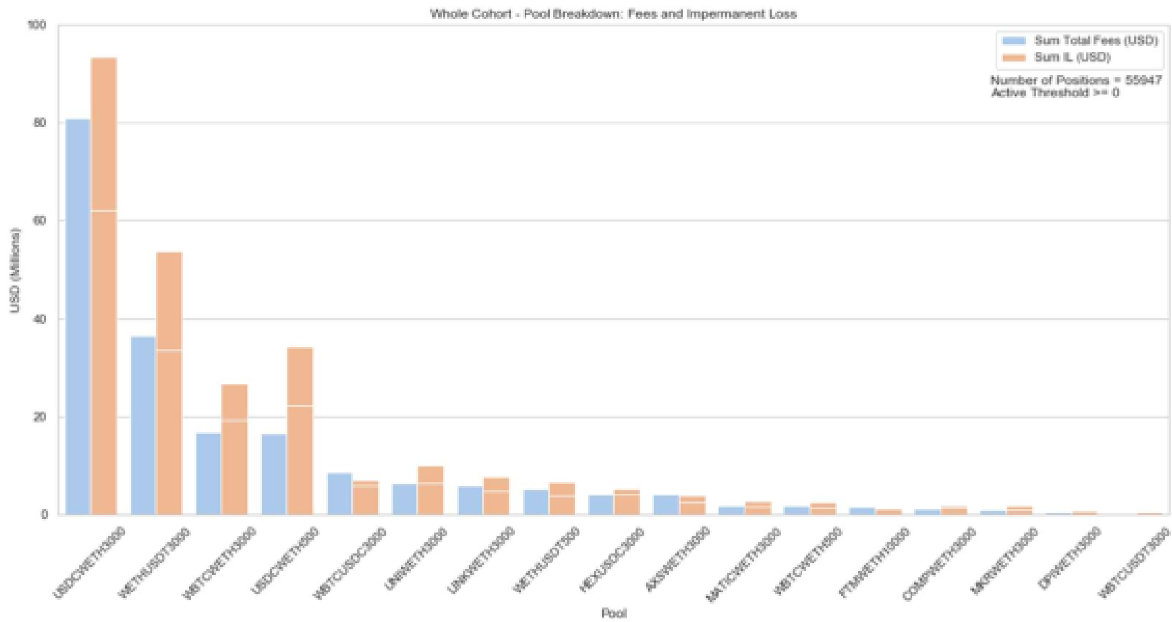
Graph 7: Impermanent loss curve

It has already been mentioned that version 3 of Uniswap introduced the concept of concentrated liquidity, which allows providers to set a range, in which the AMM operates and this creates a distinction in the understanding of IL until now. (Loesch, S., Hindman, N., Welch, N., & Richardson, M. B.; 2021) define three different kinds of IL, *minimum*, *out-of-range*, and *actual*. The first is the IL that is incurred while the asset is within the defined price range. The second is the kind that the provider suffers when the price is outside the range and is simply the difference in performance between the initial HODL position, and the frozen position at the end of the range. While minimum IL is unavoidable, out-of-range IL can be avoided if liquidity providers withdraw their funds when price steers out of their desired range. Finally, actual IL is consisted of the sum of the other two.

The study has a second part, which deals with real examples of Uniswap pools. Data collected for the analysis includes TVL and Average TVL, fees and trade volume statistics, as well as unique addresses and positions per pool. The data is then used to calculate the ROIs of different LPs in Uniswap's platform. When fee revenue is compared with the calculated IL of each pool, the researchers are left with some interesting results. In aggregate, over all the pools presented, the fees are estimated at $ 199 million whilst the IL at $ 260 million or 130% of the fees earned. This means that Uniswap providers in those pools would have been better off by $ 61 million had they simply kept their funds in their wallets. Only two pools' fees are higher than the incurred IL, namely AXSWETH3000 with $ 4 million worth of fees and IL in the area of $ 3,75 million, leaving a modest profit of $ 0,25 million. Similarly, FMTWETH10000 where the fees are $ 1,6 million and the IL is $ 1,1 million, yielding a profit of $ 0,5 million.



Graph 8: LP fee revenue vs IL (actual)



Graph 9: LP fee revenue vs IL (separated into minimum and out-of-range)

There is however a possibility of improving the first image by exiting the position after the price is out of the desired range. According to the distinction made previously, we can see that often the minimum (or unavoidable) IL does not match fee revenue in a lot of cases. This is visible in the graph above with the thin white line that separates minimum from out-of-range IL.

Overall, the findings indicate that even when accounting for price increase returns along with fees, approximately 49,5% of liquidity providers have negative net returns on their investment. The research goes on to compare even more statistic magnitudes like fees vs IL by time active and others while in the first part, it also presents more complex mathematical calculations for V3 AMM portfolio values. However, we shall not delve deeper into those as it is outside of our thesis' scope to describe this study in its completion. We merely wanted to allude to the fact that on-chain analysis has evolved further from just incorporating on-chain metrics to traditional models, as was the case in 4.2.1., and has become an inescapable reality in new and dynamic fields of study, which are on-chain by default, such as that of DeFi. For our final example, we will present a merge of traditional financial strategies with DeFi, in the same application of IL in concentrated liquidity provision AMMs.

This approach, for the most part, follows theoretical mathematics. (Deng, J., Zong, H., & Wang, Y.; 2023) define IL as:

$$\mathrm{IL} = \mathrm{Y_t} - \mathrm{Y_0} + (\mathrm{X_t} - \mathrm{X_t}) \cdot \mathrm{P_t}$$

Impermanent loss here is not considered as the difference between the entry and exit values of the position but rather as the cost of repurchasing the initial deposits when exiting the pool. Continuing, liquidity provision is tracked separately for both sides of pool price P_0, where for the "right" side we have $\mathrm{P_0} \leq \mathrm{P_l} \leq \mathrm{P_u}$ resembling the ask prices in a limit order book. For the "left" side we get $\mathrm{S_l} \leq \mathrm{S_u} \leq \mathrm{P_0}$ respectively. After a series of mathematical calculations, which we're going to emit to not get side-tracked from our original purpose, they arrive to the following proposition for the IL per liquidity (UIL):

$$\mathrm{E[UIL^R]} = -\frac{1}{2}\int_{\mathrm{P_l}}^{\mathrm{P_u}} \mathrm{K}^{-\frac{3}{2}}\mathbf{C}(\mathrm{K})\mathrm{dK}$$

$$\mathrm{E[UIL^L]} = -\frac{1}{2}\int_{S_\mathrm{l}}^{\mathrm{S_u}} \mathrm{K}^{-\frac{3}{2}}\mathbf{P}(\mathrm{K})\mathrm{dK}$$

Here, **C**(K) and **P**(K) are European call and put option prices with maturity t and strike price K. The authors refer to the above relationships as static replications of IL. They consider IL an option-like instrument that cannot easily be hedged by underlying assets and with the term "static replication" they describe the ability of the liquidity provider to buy a combination of calls or puts at inception and hold the position statically until removing liquidity from the pool, thus lowering transaction and rebalancing costs. Conversely, this generates the implication that providers can trade options in more liquid CEXs to hedge against IL through a group of call or put options with strike prices supported in the liquidity provision interval.

The thesis continues with a numerical estimation of a hypothetical position, using the Monte Carlo method under the assumption that the pool price is driven by the Heston process. The simulation results yield highly accurate approximations for both right and left side impermanent losses, if there are enough traded option strikes, and the error ratios were estimated at 0,01% for the right side and 0,001% for the left side IL. Finally, they tested the hypothesis on empirical data of BTC options traded on Deribit from 1/1/2020 to 31/12/2020. Supposedly, each day the liquidity provider can deposit funds in Uniswap's BTC – USDC pool from both sides of the current price end deplete liquidity until time *T* (one or two weeks). In the meantime, the liquidity provider also longs a combination of calls (or puts) with strikes in $[P_0, u \cdot P_0]$ (or $[d \cdot P_0, P_0]$) and holds statically to maturity *T*. The results are as follows:

Table 6: Static replication with Deribit options

T(week)	Price Interval	$\mathbb{E}[UIL^R]$	Static	#Strikes
1	$u = 1.1$	−1.588	−0.991	3
	$u = 1.2$	−2.735	−2.294	7
	$u = 1.3$	−3.597	−3.253	10
2	$u = 1.1$	−1.586	−0.921	3
	$u = 1.2$	−2.726	−2.216	7
	$u = 1.3$	−3.569	−3.157	10
		$\mathbb{E}[UIL^L]$	Static	#Strikes
1	$d = 0.9$	−0.097	−0.066	3
	$d = 0.8$	−0.154	−0.122	7
	$d = 0.7$	−0.182	−0.149	10
2	$d = 0.9$	−0.101	−0.068	3
	$d = 0.8$	−0.158	−0.124	7
	$d = 0.7$	−0.187	−0.154	10

In sum, the most interesting facts noted by the researchers were the following:

- UIL on both sides increase with wider liquidity provision intervals (larger *u*, smaller *d*)
- UIL^R is more severe than UIL^L (probably because of Bitcoin's skyrocketing from \$ 4.000 to almost \$ 30.000 in 2020)

- Replication errors are considered reasonable, especially after taking into account the large volatility of the time period and the small number of strikes.
- They were unable to spot any patterns for longer duration of liquidity provision.

5. COMPARISON WITH OTHER TYPES OF ANALYSES

We have concluded our presentation of real examples of on-chain metrics analysis in a variety of different forms and applications, ranging from simpler processes in on-line editorials to more complex statistical and other approaches in academic researches. In the introductory segments of our dissertation, we mentioned that traditional finance generally relies on different methods or different types of data from crypto. Additionally, we referenced technical analysis and how it was the only school of thought followed by crypto traders and investors in the earlier days of this new field.

Following our analysis, it should be clear by now that these different ideas and approaches to financial analysis do not seem to be diametrically opposed to one another as some might believe. In this section, we‘re going to point out points of convergence between them but also point out matters of divergence and give our humble opinion on what the future of financial analysis in the crypto spaces could look like.

Although the topic of technical analysis vs fundamental factor based propositions has been one of intense debate in the community of finance, we see that lines are begging to "turn grey" with the evolution of financial instruments and scientific progress. The advent of blockchain technology and on-chain metrics has added further towards this end. The analytical attempts we examined at unit 4 clearly show the convergence of different methodologies. We saw on-chain metrics incorporated in traditional statistical models, based on fundamental factors, sometimes alongside off-chain variables. We can infer from this that on-chain data can essentially be considered as the fundamental factors for most assets based on blockchain technology. The latter becomes explicitly clear in the last two examples, where the phenomenon of impermanent loss that was examined, is fundamentally a part of the space of digital economy and analysts tried to incorporate more traditional aspects, like options trading, to make sense of it and even hedge against it. In different spirit though, in section 4.1 we visited a few metrics which are used as, or are part of, technical indicators, similar to those we so frequently see in technical analysis. For example, the MVRV Ratio is chosen by some technical analysts as an oscillator, a very common tool in this field.

However, technical analysis strictly speaking, focuses on price action and price dynamics in the past and all its indicators revolve around values and formulae that are related to price. That is not particularly the case with on-chain indicators that can be crafted using a wide variety of data, a lot of which may not even be directly related to financial magnitudes.

Hash rates, Unique Addresses and other metrics stem from the coding aspect of the blockchain and even though they are not directly related economic factors, such as prices, they often find their way in different pieces of analysis. Furthermore, it shouldn't be mistakenly viewed that we're trying to equalize on-chain analysis with "trad-fi" fundamental factor analysis either. First of all, traditional finance employs accounting categories that sometimes are not applicable to the digital economy. On the other hand, the precision and detail that some on-chain metrics can bring to the table are not easily found in relative fields of finance in the classical sense of the word. The combination of pieces of information and the complexity of analysis that can occur from the data of blockchains is unparalleled to any other sector that we've encountered so far. Moreover, the nature of these metrics creates more possibilities in their use in complex models, which employ machine learning techniques, like those of neural networks that we also discussed. In the examined examples we saw different instances of metrics that don't really have an equivalent outside of their own context, most notably in our opinion, in the case of Realized Value. A metric, with immensely high importance in the realm of on-chain analytics and DeFi, which has no corollary to the actual world and the traditional understanding of capitalization.

On-chain metrics and on-chain indicators, in our view, bridge the gap among different methodologies in the dynamic space of the crypto economy and DeFi. Moving forward, as the digital world evolves and merges with the physical one in all kinds of different aspects, with the economy being the leading factor, we're going to see more and more overlap and merging between categories of finance and other societal functions in the two spaces. This is cause for us to believe in the immense importance of on-chain data and all their by-products, like metrics and indicators in scientific research in the years to come. Overall, we shouldn't forget that blockchain technology in general and the cryptoshpere and DeFi in particular, are still in their early stages. The room for further progress and evolution is tremendous and more data, metrics, methodologies and analyses are going to sprout from it in the future, when there is more mainstream adoption of these innovations. This leads us to the conclusion that the understanding of on-chain data and analytical methodologies is key for the proper conduction of related research going forward in time.

6. CONCLUDING REMARKS

Summarizing the content of this thesis, we began by discussing the matter of different approaches and ideas in financial analysis. We pointed out the fact that on-chain analysis is a new entry in that regard, compared to older types of fundamental analysis and technical analysis. From then on we set out to explore the nature of on-chain metrics, describe how they came to be and how they can be exploited in scientific research. Our particular field of interest is the Ethereum network. Firstly, because the subject of Bitcoin has been covered more extensively in studies and secondly, because Ethereum is more than just a means of exchange. It has given birth to a whole new space of applications and different online functions, which provide us with the challenge and opportunity of untangling their working processes and open new possibilities after studying them.

After briefly going through the ongoing progress in the body of literature, we aimed to give an analytical background feed to the reader. We describe in length basic concepts in the general ecosystem of cryptocurrencies and more particularly, we explain the modus operandi of Ethereum. We explore its blockchain, its smart contracts and DApps and we emphasize the importance of the sector of DeFi before eventually arriving at on-chain metrics. We tried to explain in a descriptive manner what exactly on-chain metrics are and provide as wide a variety of metrics as we could, attempting to cover the topic in its whole. Of course, the totality of on-chain metrics is vast and keeps rising rapidly, so it would be impossible to present a complete list. For this purpose, we highlight a few metrics that we deem more important, in terms of their commonality in analysis usage and of the possibilities they provide in fields of research.

To showcase the importance, of which we had previously talked about, we present a number of examples of analyses, which employ on-chain data from the Ethereum network in one way or another. Our first two examples come from the popular online platform “Glassnode”, whose focus is primarily on on-chain data. We present two different kinds of analysis and see how on-chain metrics are employed in analysis and exploited by traders and analysts to estimate and understand phenomena like capital flows and market sentiment. The other examples we borrow mostly from peer reviewed researches that also make use of such data. We examine more familiar cases of statistical regression models, which are modified to employ on-chain metrics inside fundamental factors approaches. Additionally, we take a look at machine learning models and more complex types of data analysis that attempt to exploit the values and properties of data that the blockchain itself can provide. In the last couple of cases we take a deeper dive into the world of DeFi and its unique phenomena. We find out the deeper workings of liquidity pools and stumble upon options trading strategies, coming to the realization that maybe there are more things in common with traditional finance than we previously realized and that the two sectors have the possibility to "communicate" with each other by hedging the risks of one using instruments from the other.

Finally, we summarize the commonalities among different approaches that we spotted during our analysis but also emphasize the fact that notable differences do exist. We end by expressing our opinion that there is a lot of room for the additional development of on-chain analysis methods with a lot of possible overlap between different schools of thought.

7. REFERENCES

Akgul, A., Şahin, E. E., & Şenol, F. Y. (2022). Blockchain-based Cryptocurrency Price Prediction with Chaos Theory, Onchain Analysis, Sentiment Analysis and Fundamental-Technical Analysis. *Chaos Theory and Applications*, *4*(3), 157–168.

Bakhtiar, T., Xiaojun, L., & Adelopo, I. (2023). The impact of fundamental factors and sentiments on the valuation of cryptocurrencies. *Blockchain: Research and Applications*, 100154.

Checkmate. (2023, September 13). The Foundational On-chain Metric: The Realized Cap. *Glassnode Insights*.

Deng, J., Zong, H., & Wang, Y. (2023). Static replication of impermanent loss for concentrated liquidity provision in decentralised markets. *Operations Research Letters*, *51*(3), 206–211.

Gunay, S., & Kaskaloglu, K. (2022). Does utilizing smart contracts induce a financial connectedness between Ethereum and non-fungible tokens? *Research in International Business and Finance*, *63*.

Heeg, K., & Schultze-Kraft, R. (2020, April 3). Introducing Account-Based On-Chain Metrics for Bitcoin and Ethereum. *Glassnode Insights*.

Jagannath, N., Barbulescu, T., Sallam, K. M., Elgendi, I., McGrath, B., Jamalipour, A., Abdel-Basset, M., & Munasinghe, K. (2021). An On-Chain Analysis-Based Approach to Predict Ethereum Prices. *IEEE Access*, *9*, 167972–167989.

Kharysh, O. (2022). *Crypto Token Price Prediction Based on Fundamental Blockchain Data*.

Kohn, A. (2023, September 4). Tracking Sentiment via Liquidity Pools. *Glassnode Insights*.

Lambert, G. (2021, June 23). Uniswap V3 LP Tokens as Perpetual Put and Call Options. *Medium*.

Loesch, S., Hindman, N., Welch, N., & Richardson, M. B. (2021). *Impermanent Loss in Uniswap v3*.

Mahmudov, M., & Puell, D. (2018, October 2). Bitcoin Market-Value-to-Realized-Value (MVRV) Ratio. *Medium*.

Said, A., Janjua, M. U., Hassan, S. U., Muzammal, Z., Saleem, T., Thaipisutikul, T., Tuarob, S., & Nawaz, R. (2021). Detailed analysis of Ethereum network on transaction behavior, community structure and link prediction. *PeerJ Computer Science*, *7*, 21–26.

Samreen, N. F., & Alalfi, M. H. (2023). An empirical study on the complexity, security and maintainability of Ethereum-based decentralized applications (DApps). *Blockchain: Research and Applications*, *4*(2).

Shah, K., Lathiya, D., Lukhi, N., Parmar, K., & Sanghvi, H. (2023). A systematic review of decentralized finance protocols. *International Journal of Intelligent Networks* (Vol. 4, pp. 171–181).

Stober, A., & Sandner, P. (2020). *Using On-Chain and Market Metrics to Analyze the Value of Crypto Assets*

Urquhart, A. (2022). Under the hood of the Ethereum blockchain. *Finance Research Letters*, *47*.