Essays on Cryptocurrency

by

Guangye Cao

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2023

Doctoral Committee:

Professor John Leahy, Chair
Professor Christopher House
Assistant Professor Pablo Ottonello
Professor Uday Rajan

Guangye Cao
guangye@umich.edu
ORCID iD: 0000-0001-7427-0802

DEDICATION

To my family.

ACKNOWLEDGMENTS

First, I would like to thank my advisors. Without Prof. Pablo Ottonello, the initial version of the model would not have come into existence. Without Professors John Leahy, Uday Rajan, and Chris House, the model's evolution would have stalled. Second, I would like to thank Marvin Ammori and Jonah Erlich, for their time spent answering my questions about token issuance in practice. Third, I would like to thank Satoshi Nakamoto for inventing the blockchain. Last but not least, I would like to thank my parents and friends for their support.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

APPENDIX

ABSTRACT

This dissertation consists of three essays. The first essay provides background on blockchain, cryptocurrency, and venture capital. It will explain the evolution of token distribution models, regulatory concerns, and the industry adoption of the technology. The second essay presents a model of startup financing that reflects regulatory concerns of the first ess ay. It develops a three-period model that compares token financing with traditional VC equity financing, where the key difference between the two is that tokens can be sold earlier than equity, which allows them to meet the liquidity needs of investors. The third essay combines token financial data and onchain transaction data from the Ethereum blockchain, to study the relationship between token liquidity, returns, and onchain market maker inventory.

CHAPTER 1

Overview of Cryptocurrency and Startup Financing

1.1 Introduction

The purpose of this chapter is to provide the reader with the background needed to motivate and understand the following chapters. I first explain how the Bitcoin and Ethereum blockchains maintain tamper-proof records of transactions without any central authority. The transaction records of the Ethereum blockchain will be analyzed in Chapter 3. I then describe two ways a blockchain startup can raise funds from venture capital investors. The first is the traditional method of issuing equity. The second way, made possible by blockchain technology, is to issue cryptocurrency that can be redeemed for the digital good produced by the startup at some future date. The difference between these two ways will be the focus of the model in Chapter 2.

1.2 Blockchain and Cryptocurrency

All currencies need ways to control supply and to enforce security properties that prevent dishonest behaviors such as counterfeiting and double-spending. Fiat currencies have physical anti-counterfeit features and rely on the central bank to control supply. A cryptocurrency is a digital currency that is not controlled by any central authority, instead, it uses cryptography and mechanism design to enforce security measures to prevent people from altering its transaction records. The transactions records are stored on a public blockchain, which is a permissionless database

with a protocol that allows users to reach consensus about its state in a decentralized manner. Permissionless means there is no restriction on who can participate.[1] Protocol means a set of rules that can be implemented by code. Decentralized means there is no central authority. The data is organized in blocks and each block is linked to the previous one, forming a chain of blocks, hence the name blockchain (Narayanan et al. (2016)).[2]

The first blockchain, Bitcoin, was invented by a pseudonymous entity named Satoshi Nakamoto, whose identity remains unknown. The cryptographic building blocks of blockchains are hash functions and digital signatures. A hash function is a mathematical function with three properties: i) its inputs can be of any size, ii) its output has a fixed size (eg: 256 bits), iii) the output can be computed in a reasonable amount of time. An example of a simple hash function is modulo hash $H(k) = k \bmod m$, where k is the input, and m is a parameter. $H(k)$ is the remainder of k divided by m. Blockchains use cryptographic hash functions, which satisfy additional properties as described in the appendix. Bitcoin uses the SHA-256 cryptography hash function, which converts all inputs into 256-bit outputs. For example, the SHA-256 hash of "apple pie" is 10ef487e48df3a7dabf54101660b74f1f1f3d5b9dc5400decda2d01099c4ccaa.

?

Each block consists of data and a hash pointer that links to the previous block. Data can be any information that is stored and encrypted on the blockchain, for example, a list of Bitcoin transactions or even code. A pointer contains the location where an object is stored. A hash pointer contains both the location and the hash of the entire previous block. The pointer shows where to find the previous block, and the hash is used to verify whether that block has been altered. As shown in figure 1.1, block 3, the latest block, contains the hash of everything in block 2, including block 2's data, the hash of block 1, and the pointer to block 1. If a hacker alters the content of block 2, then the new hash of block 2 would no longer match its old hash stored on block 3. Therefore, anyone with a copy of block 3 can verify whether the data has been tampered with. Even if the

[1] This dissertation discusses public blockchains only. There are also private blockchains that are not permissionless, i.e., access is limited to specific users. For more information on private blockchains, see Jayachandran (2017).

[2] Section 1.1 is based on the book by Narayanan et al. (2016)

hacker can modify the hash of block 2 saved in block 3, the hash of the latest block (block 3), if it had been stored separately, can be used to verify the authenticity of the entire chain.

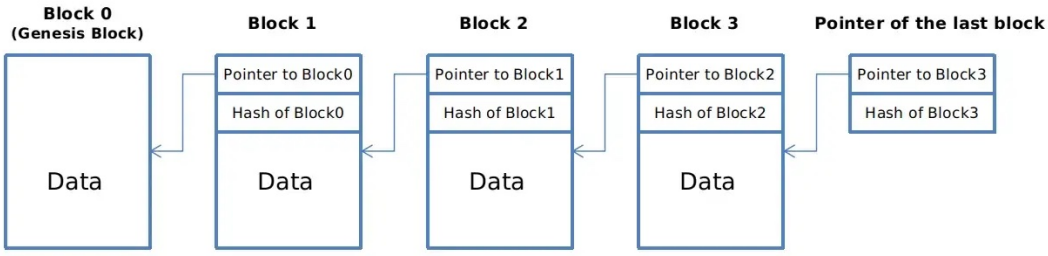

Figure 1.1: Chain of Blocks, Zhao (2018)

The second cryptographic primitive, digital signature, is an algorithm that generates a pair of keys, a public key and a secret key.[3] A public key can be considered as an identity on the blockchain. The hash of the a public key is called an $address$ and allows one to receive Bitcoins, as anyone can send Bitcoins *to* an address. To send Bitcoins *from* an address, one would need to know the secret key, which functions like a password. Because anyone can pseudonymously create an unlimited number of public and secret key pairs, it is difficult to connect addresses to their owners' actual identifies. The keys cannot be modified, so cryptocurrency owners should store their secret keys securely.

The blocks of information are stored on nodes, which are computers that each has its own copy of the entire blockchain. The nodes can send and receive information, such as new bitcoin transactions, to and from each other. Some nodes may be faulty or adversarial, for example, a node might try to alter the data stored in a previous block. Without a central authority, a *distributed consensus protocol* is needed for the nodes to reach an agreement about the state of the blockchain, i.e., which transactions have occurred and in what order. The consensus protocol needs to satisfy two properties: i) all honest nodes must reach an agreement about the state, ii) the state must have been generated by an honest node. Computer scientists have proven that when one third of the nodes are adversarial, it is impossible to reach an agreement that satisfies both properties (Lamport

[3]The properties of digital signatures can be found in the appendix.

et al. (1982)).[4] Since this problem cannot be solved with computer science, Satoshi Nakamoto turned to economics to engineer incentives for nodes to behave as desired.

The Bitcoin blockchain has processes for deciding i) how new information is added to the chain, and ii) which node gets to add the information. The economic incentives are block reward and transaction fees. If node A is selected to create and add the next block, as reward for its service, it would receive a block reward. The block reward consists of newly minted coins, so with the addition of each block, the total coin supply increases by the block reward. For Bitcoin, the block reward starts at 50 Bitcoins per block and reduces by half for every 210,000 blocks created, so the limit of Bitcoin supply is 21 million. If node A is adversarial, instead of adding the block to the end of the chain, it can try to alter the chain by creating a branch starting at an earlier block, that is, add block n to block $n - j$ where n is the block being created, and $j > 1$. Figure 1.2 gives an example of a block (V4) being added to the middle of the chain (A3) instead of to the end (A5), creating a shorter branch.

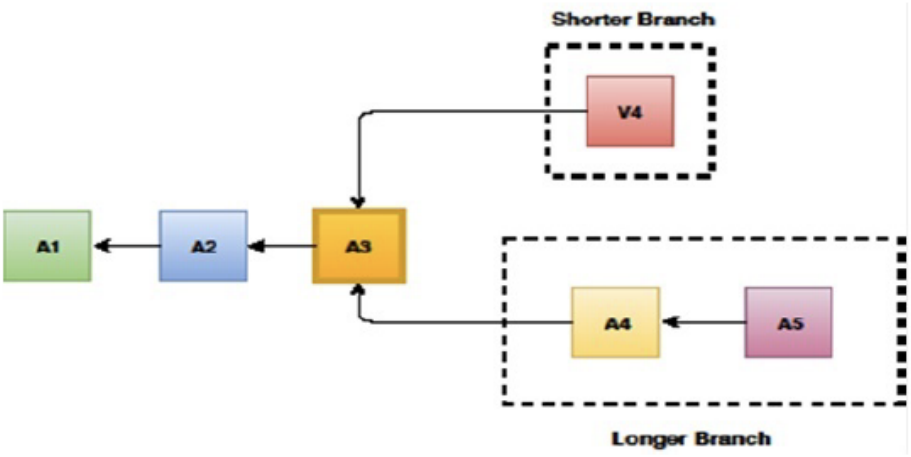

Figure 1.2: Blockchain Branches, Kaur et al. (2020)

However, by default, honest nodes append blocks to the longest chain. If an honest node B is chosen to add the following block $n + 1$, it would abandon the branch made by node A and extend the longest chain by adding A6 to A5. All the transactions recorded on A's branch would become invalid, including the transaction that sent node A the block reward; hence adversarial nodes cannot

[4]This is called the Byzantine Generals Problem.

benefit from cheating. The sender of any transaction can choose to include a voluntary transaction fee for the node that adds the block containing that transaction. Unsurprisingly, the chance that a transaction will be included in the next block increases with the generosity of the fee. Note that transaction fee consists of coins that the sender already owns, so the token supply is not expanded.

Another important innovation of Nakamoto is the process through which nodes are selected to create blocks. If every node has equal chance of being selected, an adversary would create a large number of nodes to increase its chance. The Bitcoin blockchain uses the proof-of-work process, which selects nodes in proportion to how much computing power the nodes are willing to spend. Nodes compete to solve a puzzle that requires brute force computations, so a node's probability of being picked is equal to its share of total computing power used to solve the puzzle. The puzzle is to find a value X such that when you concatenate X, the previous block's hash, and the list of transactions in the current block, the hash of this concatenation is smaller than a target number:

$$H(X||previous\ block's\ hash||tx||tx||...||tx) < target$$

where tx stands for a transaction. A node would keep trying different values for X until he gets a hash smaller than the target. Nakamoto had set a block interval parameter of ten minutes, that is, it should take ten-minutes to find X. As more computational power is dedicated to solve this puzzle, the target number is lowered to increase difficulty and maintain the ten-minute interval.

The process of finding X is called mining, and the nodes that are mining are called miners. Once a miner finds X, he includes the value in the new block, and other nodes will compute the hash of the concatenation to verify that the output is indeed below the target. This system is secure as long as none of the miners has $> 50\%$ of total computing power, that is, no miner is selected more than half of the time. Otherwise, that miner will be able to create a branch that will catch up in length to the longest chain.

Storing Bitcoins (and cryptocurrency in general) is about storing and managing one's secret keys. Typically, one would use a wallet software, which keeps track of one's balance and manages

one's keys. The software provides a simple user interface that tells the user how many coins are in the wallet. When the user wants to spend some coins, the software handles the details of which keys to use and sends the transaction messages. The software will send the information to a node controlled by the software provider, and the node will broadcast the transaction to be included in the blockchain.

While Bitcoin is an impressive breakthrough, its scripting language does not allow for loops, which restricts the usefulness of the Bitcoin blockchain to mostly payment transactions.[5] The next breakthrough is Ethereum, a blockchain that is meant to support any application imaginable, with its own native token, ether or ETH for short. Ethereum is a global, virtual computer with a Turing-complete programming language, whose state is agreed upon by the nodes in the Ethereum network.[6] Just like Bitcoin nodes, Ethereum nodes each keep a copy of the state of this computer. Unlike the Bitcoin network, which only sends coins, the Ethereum network allows nodes to send code that are executed by other nodes. These codes are called "smart contract", a term first used to describe the use of computer systems (or other automated means) to enforce contracts. For example, one can think of a vending machine as a mechanical smart contract that enforces an agreement between the user and the machine's owner regarding the purchase of a candy bar (Szabo (1997)). Anyone can create a smart contract and, for a small fee, broadcast it to the blockchain in a special transaction. The node that adds the transaction to a block will execute the contract for a fee. Each contract has its own address and can send and receive tokens.

To prevent infinite loops and limit contracts that take a long time to run, Ethereum charges "gas" to execute every operation in a smart contract. Basic operations like addition cost 1 gas, whereas writing a 254-bit word to persistent storage costs 100 gas. Every transaction also costs 21,000 gas right off the bat. Gas is paid in ETH, it is only called "gas" when being used to pay for contract execution. Each transaction can specify a "gas price", which is the number of ETH it will pay per

[5] A loop is an instruction that repeats until a specified exit condition is met. If the loop is written such that the exit condition is never met, then it is an infinite loop and will continue to execute until the program is shut off. Bitcoin does not have a way to deal of dealing with infinite loops, so it does not allow for loops at all.

[6] A Turing-complete programming language lets the user specify any functionality that is possible to program into a Turing machine, an abstract model of a computer that is believed to be capable of computing any function that can be computed at all, (Narayanan et al. (2016), p.264)

unit of gas. Miners independently decide which gas price to accept by deciding which transaction to include on the blockchain. One can think of gas as the gallons of gasoline burned by a car to drive a certain distance, and the gas price is the dollar amount that the passenger is willing to pay the driver for each gallon used. The driver will (most likely) choose the passenger willing to pay the highest gas price. Because Ethereum can only process around 15-45 transactions per second, the gas fee increases if too many transactions are submitted to be included. Figure 1.3 shows the average gas fees per transaction over time (Etherscan (2021)). To learn more about the Ethereum blockchain, see Foundation (2023).

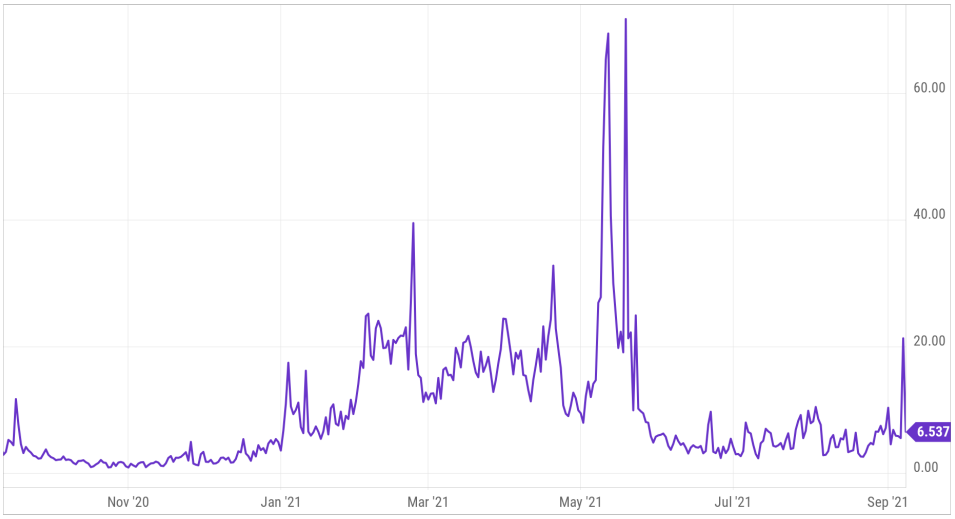

Figure 1.3: Ethereum Average Transaction Fee: USD per Tx

1.3 Startup Financing

An entrepreneur who is creating blockchain projects often need external funding from venture capital funds (VC). He can issue either equity or token in exchange for the funding. This section gives an overview of equity VC financing and token financing. For more practical information on venture capital, see Feld and Mendelson (2019).

1.3.1 Venture Capital Equity

VC investments are concentrated in a few risky high technology sectors with rapid growth opportunities (Rin et al. (2013)). Although only 0.22% of new companies created were funded by venture capital (Puri and Zarutskie (2012)), 35% of US IPOs were VC backed (Ritter (2011)). Gornall and Strebulaev (2021) studied US companies that went public after 1978, and reported that VC-backed companies accounted for a striking 92% of R&D spending. Given their role in funding innovation in general, and funding blockchain-related startups in particular, this section will describe the VC investing process.

A VC firm is structured as a limited partnership between general and limited partners. The limited partners (LP) are institutional investors who give money to the general partners (GP) to start the fund. The GPs invest in a portfolio of startups in exchange for convertible preferred equity and board seats. In return, the GPs charge LPs a management fee (2%) and keep a fraction (15 - 20%) of profit from their investments. A VC firm can have multiple funds, each with a specific focus. A fund usually has a fixed lifespan, around 7-10 years with the possibility to extend 1-3 years. Because there does not exist a liquid secondary exchange for early stage startup equity, the fund is stuck holding the equity until an exit event, that is when a startup has a IPO or is acquired. In the late 1990s, the median time from initial equity funding to IPO was approximately 3 years. After the dot com bubble, the median duration has increased to 6 - 7.5 years (figure 1.4). The median time from initial equity funding to M&A is shorter at 5 - 5.5 years (figure 1.5). Because funds cannot receive any returns for years, LPs contractually commit their investment in the fund for its entire lifespan. Therefore, VC investing is very illiquid.

Instead of giving the entrepreneur all the funding he will ever need until exit, VC financing occurs in stages (called Series). Since entrepreneur's efforts cannot be perfectly observed, staging forces them to achieve certain milestones before obtaining additional funding. Early stage companies (pre-seed to Series A) work on creating their products and proving there will be sufficient demand for them. Companies at the Series B and beyond stages need funding to scale their business. For each stage of fundraising, the entrepreneur sets a target amount that he wants to raise,

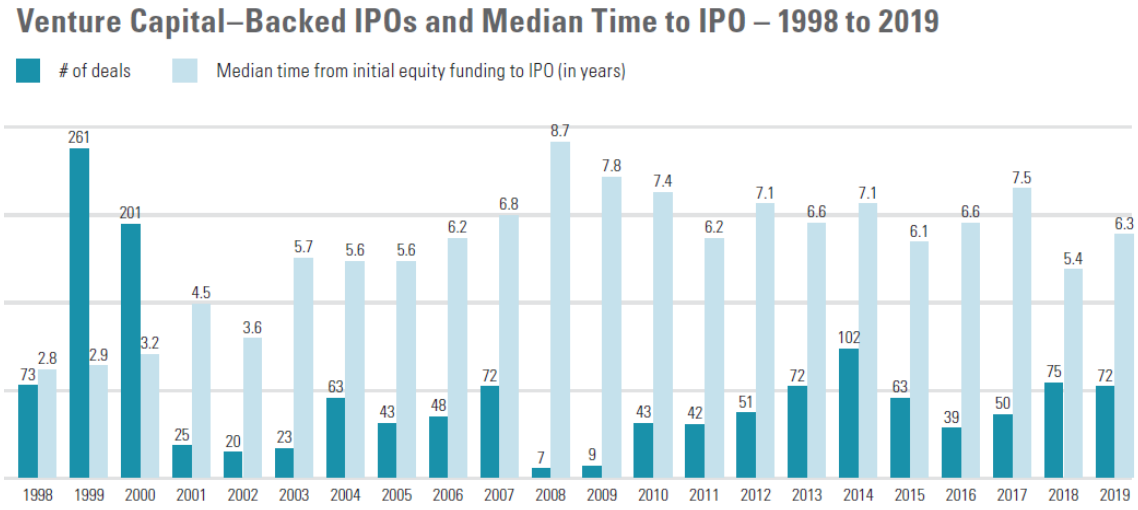

Figure 1.4: Venture Capital-Backed IPOs and Median Time to IPO - 1998 to 2019, Brasher et al. (2020)

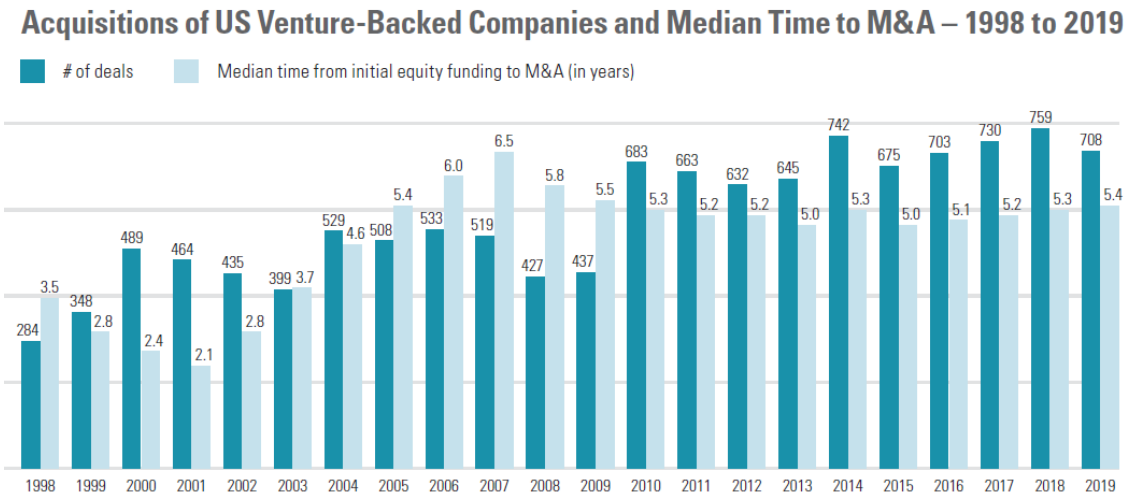

Figure 1.5: Acquisitions of US Venture-Backed Companies and Median Time to M&A, Brasher et al. (2020)

and spends weeks meeting with VCs individually to present tailored pitches. To arrange a meeting, the entrepreneur needs to demonstrate his networking ability by finding a person who is trusted by the VC to make a warm introduction, as cold emails rarely receive responses. Successful pitches will lead to a due diligence process that can take two to four weeks each. VCs often invest as a syndicate to share risk and increase deal flow, and the lead investor will negotiate the contract (called term sheet) on behalf of the syndicate. When the target amount has been secured, the entrepreneur ends the fundraise campaign to refocus on growing his business.

A critical distinction between VCs and alternatives is that only VCs provide value-adding services to their portfolio companies. Because the funds often specialize in specific industries and the best venture capitalist are former entrepreneurs themselves, top VCs have domain expertise, experience, and connections that can benefit entrepreneurs. Sahlman (1990) and Gorman and Sahlman (1989) showed that VCs spend a lot of time with their portfolio companies, sitting on the board of directors, mentoring founders, helping founders raise additional funds, and providing strategic analysis. Hsu (2004) found that high-reputation VCs offer lower valuations, and that entrepreneurs prefer low-valuation-high-reputation offers over high-valuation-low-reputation alternatives. This indicates that entrepreneurs believe it is worth accepting lower returns in exchange for working with higher quality VCs.

1.3.2 Token Financing

Blockchain entrepreneurs have an alternative to equity financing, they can instead issue tokens to investors via an initial coin offering (ICO). After the entrepreneur's product launches, the venture capitalists can resell the tokens to users or to other investors on an exchange such as Coinbase or Binance. In the early ICOs, projects[7] usually sold tokens to the public. Before starting the ICO, the founders will design a website and publish a whitepaper that explains their product and token distribution schedule. If a minimum viable product exists, the code would usually be open-sourced on Github. The founders decide the total number of tokens to be created and its allocation

[7] I use project, startup, entrepreneur, founders interchangeably

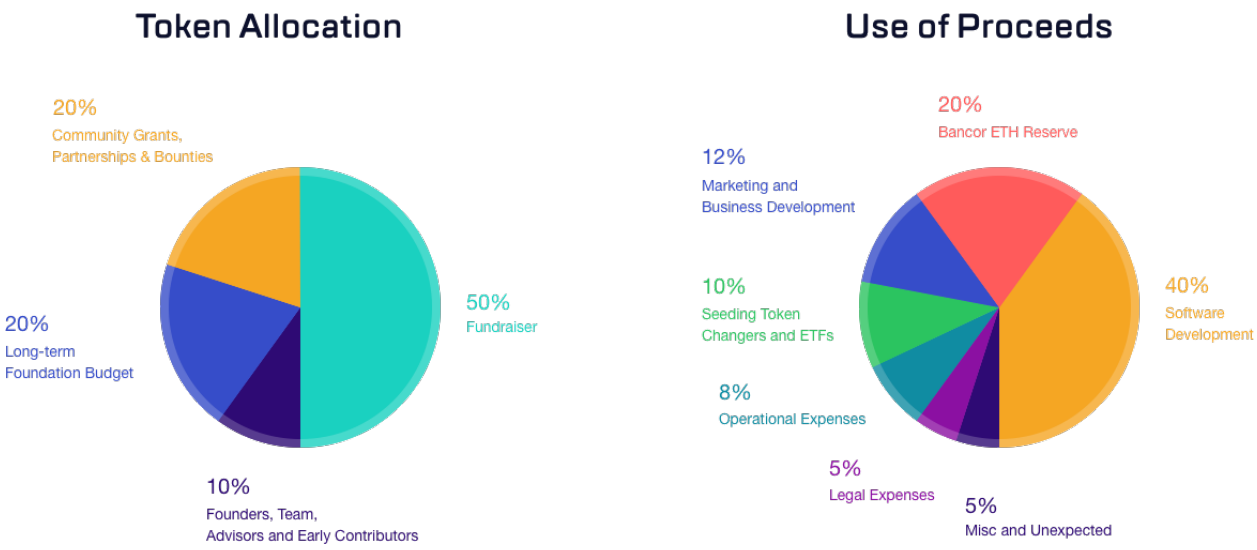

Figure 1.6: Bancor Initial Token Allocation and Use of Proceeds, Bancor (2017)

between the public (sold through the ICO), the founding team, and other stakeholders. They also announce the terms of the ICO, including the duration, participation criteria, price per token, and whether there is a cap. Figure 1.6 shows the initial allocation of Bancor, who raised around $153M by selling almost 40M BNT tokens to 10,885 buyers, at the price of 100 ETH per BNT (Bancor (2017), Higgins et al. (2017)).

The ICO attracted the attention of regulators, who questioned whether these issuances constitute unregistered securities in violation of the Securities Act of 1933. An offer of securities must be registered with the Security and Exchange Commission unless it qualifies for an exemption. A typical registration for public offering requires the filing of the S-1 registration statement, which includes audited financial statements and information about business operations, financial condition, results of operations, risk factors, and management. Pre-product blockchain startups often cannot provide some of these materials, such as the audited financial statement.

The SEC recognizes the complexity in the types of tokens and in the ways that token issuance can be structured and advertised. It stated "the analysis of whether something is a security is not static and does not strictly inhere to the instrument" (Hinman (2018)). It uses the Howey test to determine whether an asset is a security. The term security includes "investment contracts", and according to the Howey test, an "investment contract" exists when i) there is an investment of

money, ii) in a common enterprise, iii) with a reasonable expectation of profits, iv) to be derived from the efforts of others. The test analyzes not only the form and terms of the instrument, but also the "manner in which it is offered, sold, or resold (which includes secondary market sales)" (Security and Commission (2019)).[8] A direct offering of tokens to the general public would be a security because it satisfies all four prongs of the test: i) money is invested by the token buyers, ii) the startup is the common enterprise, iii) most buyers are speculators looking for profit, and iv) profit depends on the effort of the entrepreneurs.

As the Howey test is intended to protect the public, one creative approach to avoid having to face security laws is to restrict the sale of tokens to accredited investors via the use of "simple agreement to future tokens" (SAFT). A SAFT consists of two stages: i) the private sale of a promise to deliver future tokens to accredited investors, and ii) the eventual delivery of the tokens after the platform starts to operate. The initial sales constitutes an "investment contract", but because it is targeted at accredited investors instead of the public, it qualifies for an exemption under Rule 506(c) of Regulation D (Security and Commission (2022)). Proponents of SAFT are hopeful that the second stage, the actual delivery of the tokens and subsequent resale on secondary exchanges, would not be considered security. Their reasoning is that because the platform is operating, buyers on the exchange are users who purchase tokens for consumptive purposes instead of for speculation, so they do not have a "reasonable expectation of profit". Furthermore, because efforts would have already been exerted to develop the platform, future changes in token prices would no longer be "derived from the efforts of others", so the fourth prong also would not hold.

1.4 Conclusion

This chapter gives an introduction to blockchain technology and startup financing. It compares two means of financing for a blockchain startup, equity and initial coin offering, and discusses the legal pros and cons of each. It provides the background information for the second chapter, which

[8] SEC v. W.J. Howey Co., 328 U.S. 293 (1946)

develops a model of startup financing where the entrepreneur and venture capital investors decide whether to finance with equity or with token.

CHAPTER 2

Startup Financing: Token vs Equity

2.1 Introduction

Initial coin offerings (ICO) emerged as a new method of early-stage startup financing in 2013 and skyrocketed in 2017-2018. $\$11.4$ billion worth of tokens were sold in 2018, compared to seed-round venture capital issuance of $\$14.9$ billion (Pozzi (2019), Rowley (2019)). The craze subsided in 2019, when only $\$2.6$ billion was raised in the first half of the year. In an ICO, the entrepreneurs of blockchain-based startups sell cryptocurrency in exchange for bitcoin or ether[1], which they can then sell for fiat money to finance the development of their product. The cryptocurrency can have a variety of uses. It can be spent to purchase the product of the startup or the product sold on its platform, in which case it is called a utility token. In decentralized finance startups, the utility token can be deposited as collateral to mint new assets like stablecoins. Cryptocurrency can also represent a share of the profit of the startup, in which case it is called a security token.

This paper focuses on the liquidity differences between tokens and equity to answer two questions: i) when would a blockchain-based startup and venture capital investors choose to finance with tokens instead of equity; ii) what would be their rates of return for each asset?

[1] ETH or ether is the native token of the Ethereum blockchain

2.1.1 Background

A cryptocurrency is a digital asset that gives its owner access to something of value. The transaction records of these assets are maintained by a permissionless and immutable blockchain. There are multiple types of cryptocurrencies. For example, stablecoins such as DAI, Terra, and AMPL are pegged to the US dollar. If they can maintain their peg successfully, they give holders access to the stability of the USD and the ability to facilitate everyday transactions using crypto instead of fiat money. Synthetic assets, such as synthetic Tesla, is a derivative that mimics the Tesla stock, so users without access to the US stock market can have exposure to the company's stock price. Non-fungible tokens (NFTs) such as NBA Top Shot are digital collectibles whose authenticity can be verified on a blockchain.

This paper focuses on utility tokens, which give the owner access to some digital product or service on a blockchain, and are often used as currency in peer-to-peer transactions. For example, the Golem Network is a decentralized marketplace for computing power, and its GLM token is the currency used for renting spare computing power. Not all tokens are currencies for transactions, however. UNI tokens give their owners voting rights to govern the Uniswap protocol, so all stakeholders, including users, investors, and vendors, can (supposedly) influence the future of Uniswap.

The two startup financing methods, equity and initial coin offering, have been described in detail in the previous chapter. In short, equity is held until the startup goes public or is acquired, while tokens can be traded when or even before the startup's product is launched.

2.2 Literature Review

Much of the ICO literature centers around network-externality advantages of tokens for digital platform adoption, where the initial investors are also future users of the platform. Sockin and Wei (2021) investigate how tokenization resolves the conflict between platforms and users. By giving governance rights to users, tokenization prevents platforms from exploiting users, which

comes at the cost of not having an owner with equity stake who would subsidize participation to maximize network effects. Cong et al. (2021b) show that token price reflects agents' expectation of future popularity of the platform, as a permanent positive shock to platform productivity not only increases user base today, but also increases expectation of future user base and future demand for tokens, which leads agents to invest in tokens today. Other works related to network effects include Bakos and Halaburda (2019), Gryglewicz et al. (2021), and Li and Mann (2020). Another branch of ICO papers focuses on moral hazard and agency problems. For example, Chod and Lyandres (2021), and Gan et al. (2020) compare token and equity financing and argue that, token financing introduces a new agency problem - the entrepreneur underproduces the product or service, because he incurs all of the costs of production but is only entitled to a portion of the revenue. Malinova and Park (2018) argue that a variation of traditional ICO mechanism offers stronger incentives for the entrepreneur to exert effort.

To the best of my knowledge, mine is the first paper to abstract from the potentially transient differences caused by a lack of regulation, to shed light on a key difference between the two, liquidity. As ICOs have moved away from crowdfunding back towards venture capital, much of the moral hazards can be mitigated by the staged VC financing process. I therefore model startup financing where the investors are accredited investors instead of future customers. Lastly, for investors, tokens are reminiscent of the demand deposits of Diamond and Dybvig (1983) and Jacklin (1987). One difference is that the second period withdrawal (equivalent to the addiitional issuance of tokens) does not depend on the proportion of depositors who have withdrawn in the first period.

On the empirical side, Fahlenbrach and Frattaroli (2021) show that many ICO investors resell their tokens on the secondary market. Howell et al. (2020) studies the determinants of ICO success and found positive correlation with the amount of information disclosed to investors. Other notable papers include, and are not limited to: Lyandres et al. (2020), Catalini and Gans (2019), and Goldstein et al. (2022).

I build a three-period model that compares equity and tokens as means of financing for an early

stage startup building a blockchain-based product. Equity of startups is illiquid due to the lack of a secondary market, whereas tokens can be traded on cryptocurrency exchanges shortly after issuance. The model has an entrepreneur, investors (venture capital), and users. Investors have consumption needs in the middle period that can only be met with tokens. The entrepreneur's preference for tokens increases with token liquidity and investors' probability of needing to consume in the middle period. Investors willingly accept a lower return for higher liquidity.

2.3 Model

Consider a simple model in which tokens have two functions: i) as the medium of exchange to purchase the goods / services produced by the issuer (the entrepreneur), ii) as a means to raise funding for the issuer. The entrepreneur needs funding from venture capitalists to build a blockchain or a decentralized application (dApp) that will produce a digital good / service. The entrepreneur can issue either tokens or equity, where the key difference between the two assets is their liquidity. As there does not exist an exchange for trading private equity of early-stage startups, VCs normally cannot sell the equity for five to ten years, until the startup is acquired or goes public. On the other hand, tokens could be listed on public cryptocurrency exchanges within months after the private issuance. Tokens are not perfectly liquid, however, as newly listed tokens start with low trading volume due to their obscurity. Additionally, VCs are often constrained by a vesting period, so they cannot sell all the tokens at once. The model captures the liquidity advantage of tokens (relative to equity) to show that the initial price of tokens increases with liquidity. Furthermore, there is a threshold level of liquidity above which the entrepreneur can earn higher profit by selling tokens instead of selling equity.

2.3.1 Setup

There are three players: an entrepreneur, VC investors, and users; two goods: a digital good and a generic good (numeraire); three time periods: $t = \{0, 1, 2\}$, and three assets: a risk-free bond,

tokens, and equity. I use the notation $'$ for the entrepreneur, $''$ for the users, and no subscript for the investors. Assume that the initial investment I produces an exogenous stream of digital good $\{y_1', y_2'\}$, which can be interpreted as the maximum capacity or throughput of the dApp. To finance I, the cost of developing the dApp, the entrepreneur can issue equity or tokens to investors. Equity pays a dividend only in $t = 2$ and cannot be traded at all in $t = 1$. This can be interpreted as the startup needs to maintain a cash reserve by keeping the retained earnings instead of distributing them to the shareholders. Tokens can be traded in both $t = 1$ and $t = 2$; however, only a portion $\{\phi_1, \phi_2\}$ of investors' tokens can be sold in each period. One can interpret this exogenous liquidity parameter ϕ_t as the portion of tokens that can be sold with zero transaction fee, while the remaining $1 - \phi_t$ has infinite transaction fee. Alternatively,if $\phi_1 > 0$ and $\phi_2 = 1$, the parameter represents the share of investor's tokens that has become vested each period.

There are two types of investors: type a consumes only in $t = 1$, and type b consumes only in $t = 2$. In $t = 1$, investors learn of their own types. In $t = 0$, they only know that the probabilities of being types a and b are λ and $1 - \lambda$, respectively. One can think of this "consumption need" as an unexpected investment opportunity. In contrast, the entrepreneur consumes only in $t = 2$ and users consume in both $t = 1$ and $t = 2$. I further assume that users can buy both digital and generic goods, while the other players can buy only the generic good. Lastly, investors and the entrepreneur have access to all three assets, but users can only save in bonds.

The purpose of these assumptions is to isolate the effect of liquidity on token and equity pricing while capturing token's function as a medium of exchange. The assumptions that only users can buy the digital good, but they cannot save in tokens, cleanly separate the roles of users and investors. Users would not face a trade-off between keeping or selling tokens for capital gain and spending them for consumption. The assumption that the entrepreneur consumes only in $t = 2$ makes equity a feasible financing instrument for him despite its illiquidity.

I further assume that the user has perfect-substitute utility function. This means that the price of the digital good would be equal to one unit of the generic good regardless of the entrepreneur's financing choice. Lastly, when the digital and generic goods cost the same, he will consume the

digital good first. In other words, the user will buy all the digital good that is produced, within his budget constraint.

The timing is as follows. At $t = 0$, the entrepreneur either i) sells T_0 tokens at price p_0 ($p_0 T_0 = I$), or ii) create 1 share of equity priced at q and sells a fraction e of the share to investors ($eq = I$). Investors have initial wealth W, which they allocate between equity and bonds, or tokens and bond, depending on what the entrepreneur issued. The risk-free bond is completely liquid and can be traded each period costlessly. Users invest all their wealth (W'') into bonds.

At the beginning of $t = 1$, the entrepreneur launches the completed dApp to the public, and he incurs fixed cost ω to produce y_1' units of digital good. Tokens become tradeable on a cryptocurrency exchange, and investors learn of their own types. Next, asset markets clear. The entrepreneur can issue additional tokens T_1' to users and investors on the exchange at price p_1, and investors can sell up to $\phi_1 T_0$ of their previous token holding. They can also purchase bonds. Finally, users can redeem their tokens immediately for the digital good, at one token per good. If equity was issued instead, users would buy digital good with fiat money instead of tokens. Assuming the asset market clears before the goods market prevents the entrepreneur from buying back equity, as startups would generally re-invest any retained earnings back into the business.

At $t = 2$, the same process repeats. If equity was issued, the accumulated dividend would be distributed. Type b investors and entrepreneur consume the generic good.

The only variables that need to be solved for are the price and quantity of tokens and equity issued in $t = 0$ and entrepreneur's payoff from each financing method, $\{p_0, T_0, e, q, c_2'^E, c_2'^T\}$. All other variables are exogenous parameters. In the next two subsections, I solve the model with risk-neutral and risk-averse investors, separately. I will characterize the price and the required rate of return of each asset, and the entrepreneur's payoff.

2.3.2 Risk-neutral investors

2.3.2.1 Equity

User The user starts in period $t = 0$ with endowment W''. Let B''_t, c''_t, and y''_t be his risk-free bond holding, generic good, and digital good consumption, respectively. Since he can only save in bonds, $W'' = B''_0$. His optimization problems for $t \in \{1, 2\}$ are:

$$\max_{c''_1, y''_1, B''_1} (c''_1 + y''_1) + \beta \mathbb{E}_1(c''_2 + y''_2) \quad \text{subject to:} \quad B''_0 R = c''_1 + y''_1 p_1 + B''_1 \tag{2.1}$$

$$\max_{c''_2, y''_2} c''_2 + y''_2 \quad \text{subject to:} \quad B''_1 R = c''_2 + y''_2 p_2$$

Investors At $t = 0$, before knowing their own types, the investors allocate their wealth between bonds and equity to maximize expected future utility. Let Π denote the future value of profit, $\Pi = (y'_1 - \omega)R + y'_2 - \omega$.

$$\max_{B_0, e} \lambda\beta \mathbb{E}_0(\underbrace{B_0 R}_{c_1}) + (1-\lambda)\beta^2 \mathbb{E}_0(\underbrace{B_0 R^2 + e\Pi}_{c_2}) \quad \text{subject to} \quad W = B_0 + eq \tag{2.2}$$

At $t = 1$, type a investors sell all of their bonds to consume. At $t = 2$, type b investors consume their share of the startup's profit and the future value of bonds. Taking the derivative w.r.t. B_0, e, and applying $\beta = \frac{1}{R}$, the price of equity (q) is equal to the present value of expected future payoff. $\beta = \frac{1}{R}$ ensures that the players are indifferent between consuming today or saving for tomorrow.

$$q = \frac{(1-\lambda)\Pi}{R^2} \tag{2.3}$$

With $\lambda > 0$, the investors expects liquidity need in $t = 1$. While bonds can be sold to meet that need, equity cannot. To compensate, eq(2.4) shows that equity's required gross rate of return, R^E, would be greater than the risk-free rate. Investors' return is IR^E.

$$R^E = \frac{\Pi}{q} = \frac{R^2}{1-\lambda} \tag{2.4}$$

In $t = 1$, type a investors try to sell their equity to outside interests, but can only do so at a steep discount because there does not exist a liquid market for startup equity. The assumption here is that the discount price is 0, which means that type a investors' share of profits, $\lambda e \Pi$, goes to the outside interests.

Entrepreneur The entrepreneur will receive his share of accumulated profit at $t = 2$ and spend it all on the generic good (c_2'). The retained earnings from $t = 1$ accrues gross risk-free interest R. Equation 2.5 shows that the entrepreneur's payoff is equal to the future value of profit subtracted by investors' return.

$$c_{2,n}^{'E} = (1-e)\Pi = \Pi - \frac{IR^2}{1-\lambda} = \Pi - IR^E \tag{2.5}$$

2.3.2.2 Tokens

Investors

$$\max_{B_0, T_{0,n}} \lambda\beta\mathbb{E}_0 \underbrace{(B_0R + \phi_1 T_{0,n})}_{c_1} + (1-\lambda)\beta^2\mathbb{E}_0[\underbrace{(B_0R + \phi_1 T_{0,n})R + \phi_2(1-\phi_1)T_{0,n}}_{c_2}]$$

$$\text{subject to:} \quad W = B_0 + T_{0,n}p_{0,n} \quad \text{and}$$

$$T_{1,n} \geq (1-\phi_0)T_{0,n} \quad \text{(liquidity constraint)}$$

Because $p_1 = p_2 = 1$, investors will not earn returns from holding tokens from $t = 1$ to $t = 2$. They will sell as much of their token holding as they can, so the liquidity constraint is binding $T_{1,n} = (1-\phi_1)T_{0,n}$. Taking the derivative w.r.t. B_0 and $T_{0,n}$, and applying $\beta = \frac{1}{R}$, eq. 2.6 shows that the price of token ($p_{0,n}$) is equal to the present value of the fraction of tokens sold in $t = 1$, plus the present value of the additional portion sold in $t = 2$ multiplied by the probability of needing liquidity in $t = 2$. Since some tokens cannot be sold in $t = 1$ to meet the needs of type a investors,

tokens become less valuable as the need for early consumption (λ) increases. Token price increases in ϕ_1 and ϕ_2 as being able to offload tokens in each period makes the asset more valuable.

$$p_{0,n} = \frac{\phi_1}{R} + \frac{(1-\lambda)\phi_2(1-\phi_1)}{R^2} \tag{2.6}$$

In eq. 2.6, $\frac{\phi_1}{R}$ is the present value of utility from selling tokens in $t = 1$, and $\frac{(1-\lambda)\phi_2(1-\phi_1)}{R^2}$ is the utility from selling additional tokens in $t = 2$ if the investor is type b. The expected required gross rate of return by selling at $t = 1$ for $p_1 = 1$ is:

$$R_n^T = \frac{1}{p_{0,n}} = \frac{R^2}{\phi_1 R + (1-\lambda)\phi_2(1-\phi_1)} \tag{2.7}$$

Entrepreneur: The entrepreneur's payoff is the value of $t \in \{1, 2\}$ token issuance net of ω, which depends on the parameters ϕ_1 and ϕ_2. The tokens he issues in a period is equal to digital good output minus the quantity of tokens resold by investors; $T'_{1,n} = y'_1 - \phi_1 T_{0,n}$ and $T'_{2,n} = y'_2 - \phi_2(1-\phi_1)T_{0,n}$. Equation 2.3.2.2 shows that entrepreneur's payoff decreases in λ. Note that $T_{0,n} = \frac{IR^2}{\phi_1 R + (1-\lambda)\phi_2(1-\phi_1)}$ is increasing in λ, meaning as the probability of needing early liquidity increases, token price must decrease to compensate for the asset's imperfect liquidity in $t = 1$. As more tokens must be issued in $t = 0$ to finance I, fewer will be issued by the entrepreneur afterwards.

$$\begin{aligned} c_{2,n}^{'T} &= T'_{2,n} p_2 + (T'_{1,n} p_1 - \omega)R - \omega \\ &= (y_1 - \omega)R + y_2 - \omega - IR^2\left(\frac{1}{\frac{\lambda R \phi_1}{\phi_2(1-\phi_1)+\phi_1 R} + 1 - \lambda}\right) \\ &= \Pi - \frac{IR^2}{\frac{\lambda R \phi_1}{\phi_2(1-\phi_1)+\phi_1 R} + 1 - \lambda} \\ &\geq c_{2,n}^{'E} = \Pi - \frac{IR^2}{1-\lambda} \end{aligned} \tag{2.8}$$

Result 1: $\frac{\partial c_{2,n}^{'T}}{\partial \phi_1} > 0$ and $\frac{\partial c_{2,n}^{'T}}{\partial \phi_2} < 0$. *Proof*: see technical appendix.

The entrepreneur's payoff increases in ϕ_1 and decreases in ϕ_2. Higher ϕ_t makes tokens more valuable so fewer are sold in $t = 0$, which benefits the entrepreneur as he will be able to sell more new tokens in $t \in \{1, 2\}$ to meet demand for digital good. On the other hand, higher ϕ_t also allows investors to resell more tokens in $t \in \{1, 2\}$, so the entrepreneur would issue fewer new tokens. For ϕ_2, the latter effect is stronger than the former. In the special cases where $\phi_1 = 0$ and $\phi_2 > 0$ or $\phi_1 > 0$ and $\phi_2 = 0$, the number of tokens resold by investors, $\phi_1 T_0 = \phi_1 \frac{IR}{\phi_1}$ and $\phi_2 T_0 = \phi_2 \frac{IR^2}{(1-\lambda)\phi_2}$, do not depend on ϕ_t, meaning that T_0 adjusts perfectly to keep entrepreneur payoff independent of ϕ_t.

Result 2: $c_{2,n}^{'T} > c_{2,n}^{'E}$ if $\phi_1 > 0$ and $\lambda > 0$.

Equation 2.3.2.2 shows that if $\phi_1 > 0$ and $\lambda > 0$, then $\frac{\lambda R \phi_1}{\phi_2(1-\phi_1)+\phi_1 R} + 1 - \lambda > 1 - \lambda \Rightarrow \frac{IR^2}{\frac{\lambda R \phi_1}{\phi_2(1-\phi_1)+\phi_1 R}+1-\lambda} < \frac{IR^2}{1-\lambda} \Rightarrow \Pi - \frac{IR^2}{\frac{\lambda R \phi_1}{\phi_2(1-\phi_1)+\phi_1 R}+1-\lambda} \geq \Pi - \frac{IR^2}{1-\lambda}$. This means as long as there is positive need for early liquidity and tokens can partially satisfy that need, the entrepreneur is better off issuing tokens instead of equity. Even if $\phi_2 = 0$, $c_{2,n}^{'T} = \Pi - IR^2 > c_{2,n}^{'E} = \Pi - \frac{IR^2}{1-\lambda}$. Intuitively, even if type b investors cannot sell any tokens in $t = 2$, they can sell some tokens in $t = 1$ and still be able to consume in $t = 2$. The value of ϕ_2 does not matter because $\phi_1 > 0$ guarantees consumption in both period.

Result 3: Equity is the same as tokens that can be sold only in $t = 2$, ie: $\phi_1 = 0$, $\phi_2 = 1$. In this case, $p_0 = (1 - \lambda)\beta^2 = \frac{1-\lambda}{R^2}$. Investors' required rate of return and entrepreneur's payoff are the same as those with equity: $R^T = \frac{1}{p_0} = \frac{R^2}{1-\lambda} = R^E$ and $c_{2,n}^{'T} = \Pi - \frac{IR_f^2}{1-\lambda} = c_{2,n}^{'E}$.

Result 4: The risk-free bond is the same as tokens can be all sold in $t = 1$. In this case, $p_0 = \frac{1}{R}$, $R^T = R$, and $c_{2,n}^{'T} = \Pi - IR^2$.

To summarize, with risk-neutral investors, the price of tokens at $t = 0$ is the present value of utility from selling tokens in $t = 1$ plus the expected utility from selling additional token in $t = 2$ if the investor is revealed to be type b. As long as some tokens can be sold in $t = 1$, ie. $\phi_1 > 0$,

tokens guarantee that both types of consumers will be able to consume, while equity allow only type b consumers to consume.

2.3.3 Risk-averse investors

2.3.3.1 Equity Financing

Investors: With constant relative risk aversion, investors' desire to smooth consumption increases with curvature (σ) of the utility function. Equity cannot satisfy this desire at all as it gives investors no return at $t = 1$. Hence, CRRA investors value equity less than risk-neutral investors would (equation 2.10).

$$\max_{B_0,e_a} \lambda\beta\mathbb{E}_0\frac{(\overbrace{B_0R}^{c_1})^{1-\sigma}-1}{1-\sigma}+(1-\lambda)\beta^2\mathbb{E}_0\frac{(\overbrace{B_0R^2+e_a\Pi}^{c_2})^{1-\sigma}-1}{1-\sigma} \quad \text{subject to} \quad W = B_0 + e_a q \tag{2.9}$$

$$q_a = \frac{(1-\lambda)\Pi}{R^2\left[\lambda(\frac{B_0R^2+e_a\Pi}{B_0R})^\sigma+1-\lambda\right]} \leq \frac{(1-\lambda)\Pi}{R^2} = q_n \tag{2.10}$$

$$R_a^E = \frac{\Pi}{q_a} = \frac{R^2}{1-\lambda}\left[\lambda(\frac{B_0R^2+e_a\Pi}{B_0R})^\sigma+1-\lambda\right] \geq \frac{R^2}{1-\lambda} = R_n^E \tag{2.11}$$

A larger $\frac{B_0R^2+e_a\Pi}{B_0R} = \frac{c_2}{c_1}$ indicates less consumption smoothing. In equation 2.11 shows that investors require a risk premium $\lambda(\frac{B_0R^2+e_a\Pi}{B_0R})^\sigma+1-\lambda$ that increases in $\frac{c_2}{c_1}$, leaving less profit for the entrepreneur.

Entrepreneur

$$c_{2,a}^{\prime \mathrm{E}} = \Pi - \frac{IR^2}{1-\lambda}\left[\lambda(\frac{B_0R^2+e_a\Pi}{B_0R})^\sigma+1-\lambda\right] \leq c_{2,n}^{\prime \mathrm{E}} \tag{2.12}$$

2.3.3.2 Token Financing

Investor

Equation 2.13 shows that token price with CRRA investors is equal token price with risk-neutral investors adjusted for the ratio $\frac{c_2^{-\sigma}}{\lambda c_1^{-\sigma}+(1-\lambda)c_2^{-\sigma}}$, which represents period 2 marginal utility as a share

of expected marginal utility. A smaller ratio indicates less consumption smoothing with tokens.

$$p_{0,a} = \frac{\phi_1}{R} + \frac{(1-\lambda)\phi_2(1-\phi_1)}{R^2} \frac{\overbrace{c_2^{-\sigma}}^{\text{MU}_2}}{\underbrace{\lambda c_1^{-\sigma} + (1-\lambda)c_2^{-\sigma}}_{\text{expected marginal utility}}} \tag{2.13}$$
$$\leq \frac{\phi_1}{R} + \frac{(1-\lambda)\phi_2(1-\phi_1)}{R^2} = p_{0,n}$$

Entrepreneur

Result 5: The entrepreneur's payoff (consumption in $t = 2$) declines as investors become more risk-averse (as σ increases). The decline is more steep if the entrepreneur finances with equity instead of tokens. See figure 2.1 as illustration.

$$c_{2,a}^{\prime \text{T}} = \Pi - \frac{IR^2(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma)(\phi_2(1-\phi_1) + \phi_1 R)}{\phi_1 R(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma) + (1-\lambda)\phi_2(1-\phi_1)c_1^\sigma} \tag{2.14}$$
$$\leq \Pi - IR^2 \frac{\phi_2(1-\phi_1) + \phi_1 R}{\phi_1 R + (1-\lambda)\phi_2(1-\phi_1)} = c_{2,n}^{\prime \text{T}}$$

2.4 Conclusion

This simple model compares an entrepreneur's payoffs when financing by issuing equity vs tokens to venture capital. Tokens allows both types of investors to consume, whereas equity satisfies only the investors with late consumption needs. Before knowing their own types, investors are willing to accept a lower rate of return in exchange for the liquidity benefits of tokens. For the entrepreneur, tokens give higher payoff than equity, and the difference in payoff increases with investors' risk aversion.

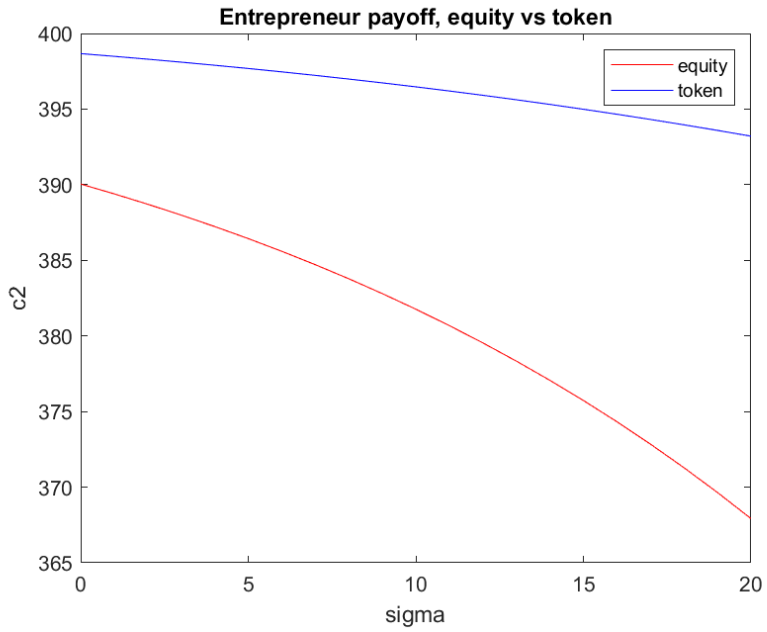

Figure 2.1: Entrepreneur Payoff, Equity vs. Token, for $\lambda = 0.1$

CHAPTER 3

Token Liquidity

3.1 Introduction

Chapter 2 showed that liquidity might play a role in token financing. This chapter investigates whether liquidity is priced in crypto markets. I study the cross-section relationship between cryptocurrency liquidity and returns. Furthermore, I build a novel dataset from onchain token transfers, define onchain market makers (OMM), and study how OMM activity and inventory correlate with returns.

I use the Amihud (2002) (il)liquidity measure, which is the daily ratio of the absolute value of token return to dollar trading volume, averaged over a week. It can be interpreted as the daily price response associated with one dollar of trading volume. Higher Amihud ratio indicates higher price impact, hence higher illiquidity. While there are better measures of liquidity, such as the bid-ask spread, these measures require transaction-level price data that are not available for the majority of cryptocurrencies. Even when available, the data is only obtainable from a limited number of exchanges, while most tokens are not traded on the included exchanges. I use the terms "Amihud ratio" and illiquidity interchangeably.

Onchain data is data recorded on a blockchain. I use data from Ethereum, the dominant blockchain for building decentralized applications (dApps). Ethereum is analogous to an operating system such as the iOS, on which applications like Uber can be deployed. A dApp can issue its own native token, which is often used as the medium of exchange to purchase digital services

through the dApp.[1] These tokens usually follow the ERC-20 standard, which defines a common set of functions and syntax for fungible tokens.[2] I study ERC-20 tokens for three reasons: i) many of the most widely-adopted and influential dApps have issued ERC-20 tokens, ii) Ethereum is the dominant blockchain for building decentralized applications, as there are more projects being built on top of Ethereum than on its competitors, iii) ERC-20 token transaction data are obtainable with relative ease from a single source, Google Bigquery.

Because Ethereum records every token transfer, the blockchain itself is a treasure trove of real-time transaction data. Onchain data is composed of sender and receiver addresses, timestamp, token identification, and quantity of token transferred. Unlike traditional startups whose business transactions are not public information, the records of a blockchain startup's activities are publicly available on the blockchain. This transaction-level data could give insights into the growth and adoption of the startup's product or service and speculation of its token. I combine onchain data with financial data from exchanges to study the behavior of onchain market makers and liquidity.

I define an onchain market maker (OMM) for token i on day t as an address that has received and sent more than 1% of token i's daily onchain transfer volume. In other words, an OMM is an address on both the buying and selling sides of the market, who is also large enough to account for a significant portion of the tokens transferred in a day.

The results show that across tokens, liquidity premium exists in the cryptocurrency market, as token return increases with illiquidity. If a token's mean-adjusted illiquidity, that is, the ratio of token illiquidity to market illiquidity, increases by 1 (i.e., 100 percentage points (p.p.)) then its weekly, non-annualized return is expected to increase by 0.088 p.p. the following week, controlling for mean-reversion, size, beta, and volatility. This is greater than the liquidity premium of stocks, which Amihud (2002) found to be 0.112 p.p. using non-annualized monthly instead of weekly returns.

[1] Tokens can have other functions, for example, giving the owner voting rights.

[2] ERC-20 stands for Ethereum Request for Comment 20. Functions can include methods such as totalSupply() or events such as Transfer(). For more detail on ERC-20 standard, please refer to: https://ethereum.org/en/developers/docs/standards/tokens/erc-20/. ERC-20 is specific to Ethereum, other blockchains have their own fungible token standards. For example, the Binance Smart Chain has BEP-20 tokens.

Furthermore, there is a negative correlation between returns and onchain market makers' contemporaneous change in inventory, meaning the OMMs provide liquidity by taking the opposite side of the market to accommodate buying and selling pressure. A 10 p.p. increase in OMMs' net token purchase relative to total token transfers is correlated with a 0.62 p.p. decrease in contemporaneous weekly returns.

This chapter unfolds as follows. Section 2 provides a brief literature review. Section 3 provides a crypto market overview. Section 4 presents a few facts about onchain market makers. Section 5 presents cross-section estimates of token returns as a function of the Amihud ratio and other variables. Section 6 concludes.

3.2 Literature Review

This paper is related to the literature on liquidity and asset pricing. Acharya and Pedersen (2005) presents a model of liquidity-adjusted CAPM where a persistent negative shock to a security's liquidity results in low contemporaneous returns and high predicted future returns. Empirically, the relationship between liquidity and expected return has been studied for a variety of assets. Amihud (2002) found negative relationship between future stock returns and liquidity, Lin et al. (2011) and Mancini et al. (2013) found similar dynamics for bonds and foreign exchange. Also using the Amihud ratio to study illiquidity in cryptocurrency, Zhang and Yi (2023) found evidence of liquidity premium that is significant at 5%, while controlling for standard risk variables such as size and momentum. On the other hand, again using the Amihud ratio, Liu et al. (2022) found that the return difference between the most and least liquid portfolios are only significant at 10%.[3] One contribution of the current paper is the addition of novel onchain variables, such as onchain market maker activity and the share of addresses.

This paper also contributes to the literature on market maker inventory, because net token purchase by OMMs is equivalent to changes in market maker's inventory. Comerton-Forde et al.

[3]Liu et al. (2022) kept stablecoins in their data. Stablecoins are tokens whose values are pegged 1:1 to USD, so their returns are intended to be 0.

(2010) shows that market maker inventory and revenue explains time-variation in liquidity. Hendershott and Seasholes (2007) shows that market maker inventory is negatively correlated with contemporaneous returns, which is consistent with market makers temporarily accommodate buying and selling pressure. This paper observes similar relationship in the crypto market, where OMMs are net token buyers when contemporaneous returns decreases.

3.3 Crypto Market Overview

I scraped from coinmarketcap.com the financial data of the 2000 tokens with the highest market capitalization as of March 1st, 2021. The cryptocurrency market is highly segregated, with each token often trading on multiple exchanges. Coinmarketcap is an aggregator that collects data from exchanges' APIs to compute the weighted average price for each token. Issues inherent to aggregated data include: i) exchanges that do not have APIs are excluded, and ii) crypto exchanges are notorious for wash trading, so the reported trading volume may not be accurate (Cong et al. (2021a)). [4]

Figure 3.1 shows the mean, median, standard deviation of the monthly returns of all tokens each month. Both mean and median return vary greatly between months, and the standard deviation shows that there is large variation in returns between tokens within the same month. The blue line (right axis) represents the number of tokens available each month. Token-days with less than $1000 in trading volume are dropped before aggregating to monthly frequency, leaving around 1600 tokens in the data.

[4]Wash trading occurs when an exchange reports trading volume that did not exist, or inflates trading volume by buying and selling tokens itself.

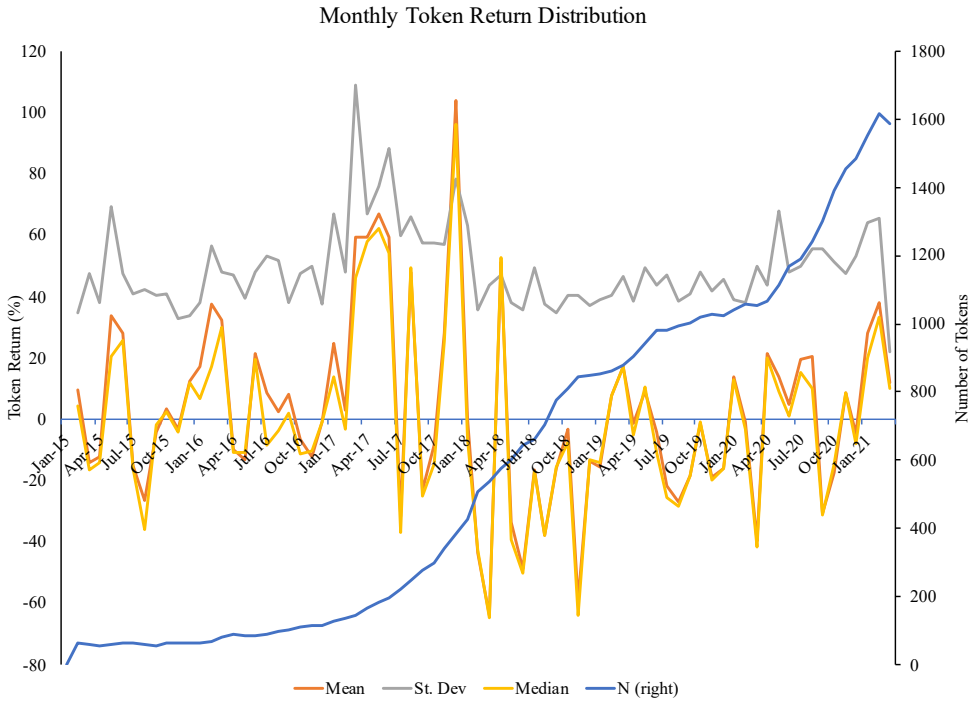

Figure 3.1: Monthy Token Return Distribution, and Number of Tokens

Figure 3.2 shows the quartiles of returns across all available tokens each month. Returns are highly skewed in some months.

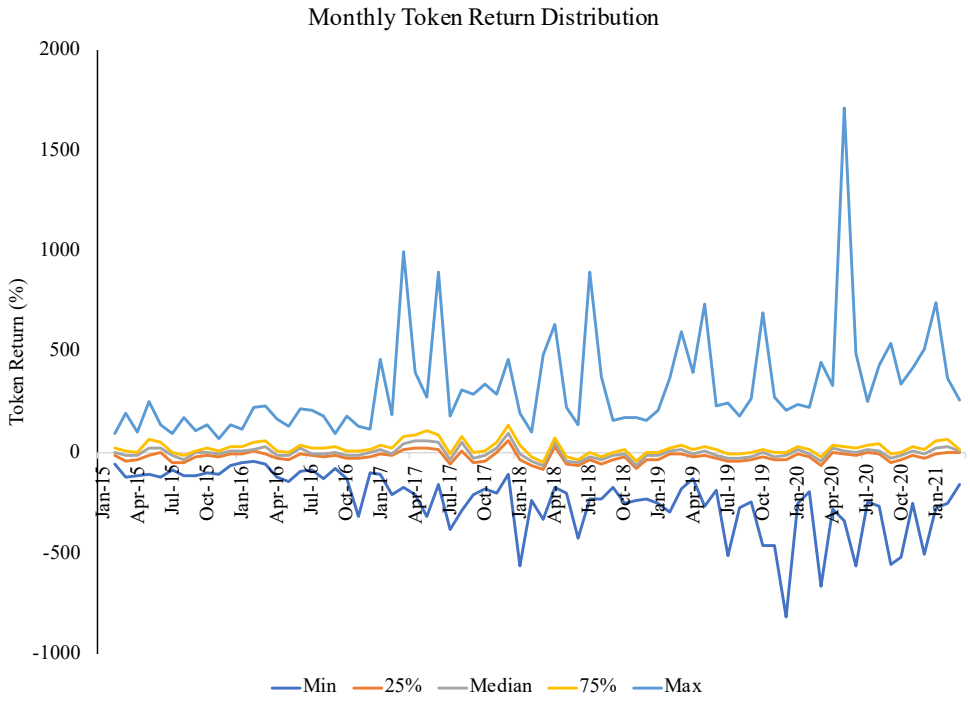

Figure 3.2: Monthly Token Return Distribution, Percentiles

Next, I delete stablecoins, which are tokens pegged to fiat currencies, and construct a crypto market price index by computing the average token price weighted by market capitalization. Figure 3.3 displays price index (green) with the number of tokens (blue). The market price index will be used to estimate each token's beta in the regression analysis.

Figure 3.4 shows the number of ERC-20 tokens in my data. There are tens of thousands of ERC-20 tokens in existence, but most of them are not traded, and only around 1200 belong to the top 2000 by market capitalization.

Figur 3.5 shows the weekly return of the crypto market price in from figure 1. Return is calculated as log difference. The returns are highly volatile, varying between -40% and 40%. The fluctuation is also evident in figure 3.6, the crypto market 30-day return standard deviation.

Figures 3.7 displays the market capitalization of Bitcoin, Ether (the native token of Ethereum

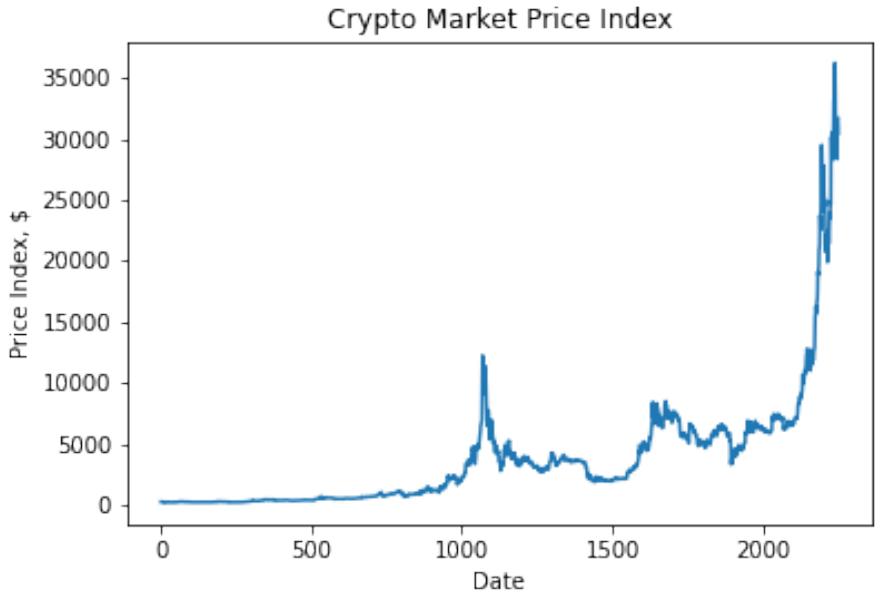

Figure 3.3: Crypto Market Price Index

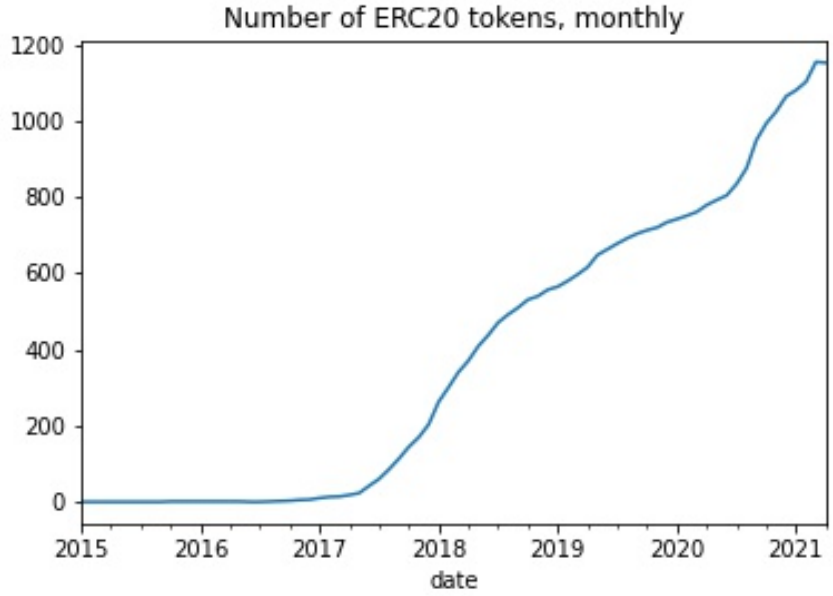

Figure 3.4: Number of ERC20 Tokens, Monthly

blockchain), sum of all ERC-20 tokens, and others. ERC-20 tokens still accounts for less than 20% of total market capitalization, while Bitcoin continues to dominate with 60%.[5]

[5]Bitcoin dominance has further dropped to below 50% since March 2021

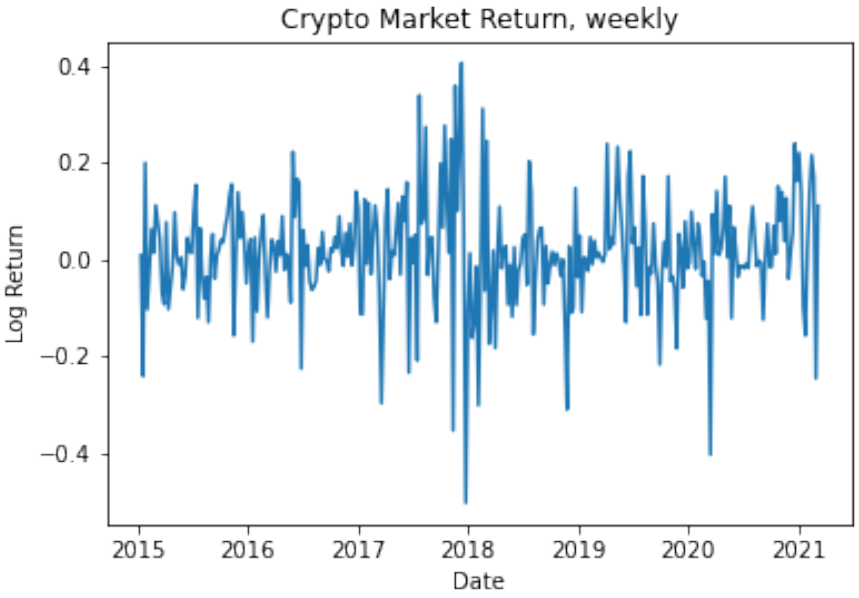

Figure 3.5: Crypto Market Weekly Return

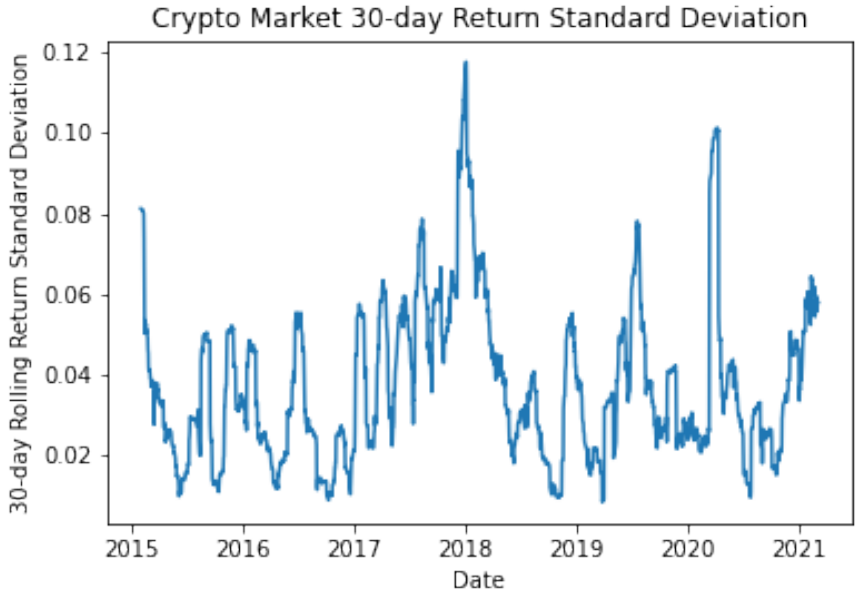

Figure 3.6: Crypto Market Return 30-day Volatility

3.4 Facts about Onchain Market Makers:

As mentioned in the introduction, an onchain market maker (OMM) for token i on day t is an address that has received and sent more than 1% of token i's daily onchain transfer volume. In

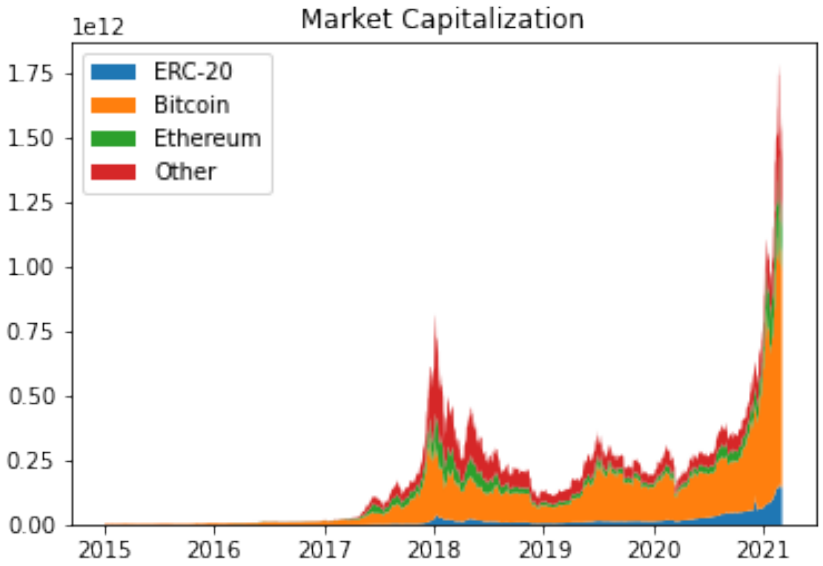

Figure 3.7: Market Capitalization, USD

other words, an OMM is an address on both the buying and selling sides of the market, who is also large enough to account for a significant portion of the tokens transferred in a day. OMMs are determined on token-day basis, so a OMM for token i on day t might no longer be OMM on day $t + 1$, either because it is no longer on both sides of the market or because it transferred less than 1% of i's volume. Please note that onchain *transfer volume*, which is the amount of token i transferred in a day, is different from the dollar *trading volume* that occur on exchanges. Trades on centralized exchanges such as FTX often only appear in the exchanges' own records, but not on the blockchain. Token transfers on the blockchain are not counted in trading volume unless the transfers are to or from an exchange.

Below are a few facts about onchain market maker activity.

1. OMMs market make a token for 11.3 days per month on average. It is not clear what OMM do the rest of the month, perhaps trading other tokens.

2. OMMs account for a large share of transfer volume relative to their numbers. On average, OMMs make up 30.2% and 28.2% of all receiving and selling addresses, respectively. They

receive (or buy) and send (or sell) 56% and 50.9% of total trading volume.

3. OMMs are net buyers when returns are low. Let net purchase share$_{i,t}$ be OMM's net purchase of token i on day t as $\%$ of total transfer volume. $r_{i,t}$ is token i's return on day t in percent, and γ_i is token fixed effects. The fixed effects estimation shows that when OMMs' net purchase share increases by 1 percentage point, same-day return of the token decreases by 0.025 p.p., after controlling for the aggregate crypto market, lag return, and token fixed effect.

$$\begin{aligned} r_{i,t} = &-0.34^{**} - 0.24^{**} r_{i,t-1} + 0.64^{**} \text{crypto market index return}_t \\ &- 0.025^{**} \text{net purchase share}_{i,t} + \gamma_i + \epsilon_{i,t} \end{aligned} \tag{3.1}$$

3.5 Empirical Methodology

Following the Fama and MacBeth (1973) regression in Amihud (2002), I run a cross-section model for each week. I regress each week's token return on the first lags of financial variables and contemporaneous onchain variables. For each coefficient, I then average it over all weeks, and compute t-statistics to test whether the average of each coefficient is significantly different from 0.

$$R_{iw} = \alpha_w + \beta_{0w} R_{i,w-1} + \sum_{j=1}^{n} \beta_{jw} X_{ji,w-1} + \sum_{j=n+1}^{J} \beta_{jw} X_{ji,w} + \epsilon_{iw} \tag{3.2}$$

R_{iw} is token i's return in week w in percentage. $X_{0i,w-1}...X_{ni,w-1}$ are the lags of weekly averages of the daily values of financial variables, and $X_{n+1i,w}...X_{Ji,w}$ are the contemporaneous onchain variables, including:

1. *Mean-adjusted Amihud ratio*: since average liquidity varies considerably over the weeks, I calculate the mean-adjusted Amihud ratio by dividing each token's weekly average Amihud ratio by the market's weekly Amihud ratio. Let token i's weekly illiquidity be Amihud$_{i,w} = \frac{1}{T}\sum_{t=1}^{T}$ Amihud$_{i,t}$, where T is the number of days in week w for which token i's Amihud ratio is available. Let the weekly market illiquidity be the simple average of token illiquidity:

$\text{Amihud}_{market,w} = \frac{1}{N_w} \sum_{i=1}^{N_w} \text{Amihud}_{i,w}$, where N_w is the number of tokens available in week w. The mean-adjusted Amihud is:

$$\text{AmihudMA}_{i,w} = \frac{\text{Amihud}_{i,w}}{\text{Amihud}_{market,w}}$$

2. *Log size*: natural logarithm of a token's market capitalization.

3. *Beta*: token i's quarterly comovement with market returns. Token i's quarterly beta is estimated from:

$$R_{iq} = \alpha_{iq} + \text{beta}_{iq} RM_q + \epsilon_{iq} \tag{3.3}$$

where R_{iq} is token i's return in quarter q, and RM_q is the same period's market return.

4. *SDRET*: quarterly standard deviation of daily return of token i in quarter q. By the asset pricing models of Levy (1978) and Merton (1987), SDRET is risk that is priced because investors' portfolio are constrained and therefore not well diversified. Stoll (1978) suggested that stock illiquidity is positively related to the stock's risk since the bid-ask spread set by a risk-averse market maker is increasing in the stock's risk. Constantinides (1986) proposed that stock variance positively affects the return that investors require on the stock, since it imposes higher trading costs on them due to the need to re-balance their portfolios more frequently.

5. *Buy (sell)-share*: the amount of token i bought (sold) by i's OMMs as fraction of daily transfer volume. This is an indication of the level of activity by OMMs. A larger fraction means more market makers are trading token i, making it more liquid.

6. *Whales*: the $\%$ of token i's OMM who are also whales. Whales are addresses that hold more than 1% of a token's supply.

7. *Net purchase share*: token i's OMM net token purchase divided by total transfer volume.

8. *Positive address share*: an address for token i on day t refers to an onchain address that has received or sent token i at any time in the past and present, i.e., until and including t. Positive address share is the fraction of addresses that has positive balance of token i on day t.

9. *Zero address share*: the fraction of addresses for token i that no longer own the token on day t, ie., their balance of token i is zero on day t. Zero balance addresses could indicate that the address owner has abandoned the token. An increase in zero address share could mean a loss of interest in token i by the market.

10. *New address share*: the fraction of addresses for token i that sent or received token i for the first time on day t.

To be included in the data, observations must meet the following criteria:

- Tokens must follow the ERC-20 format. Non-ERC-20 tokens, such as non-fungible tokens (NFTs) are eliminated because each unit can represent a different asset.
- Tokens whose value is derived from other assets are eliminated. An example would be stablecoins that are pegged to fiat currencies.
- Tokens with fewer than 180-days of financial and onchain data are eliminated.
- For each day, observations within the top and bottom 0.5% of Amihud ratios are deleted.
- Weeks that have fewer than 100 tokens with available onchain and financial data are deleted.

892 tokens meet all of the above criteria. The data spans from the first week of 2018 to the eighth week of 2021, for a total of 164 weeks. Table 3.1 summarizes the variables of interest. Weekly token returns exhibits large variation, ranging from -695% to 884%, with mean -0.58% and median -0.78%. Token illiquidity is highly skewed, with median 0.000062, i.e., $\$1$ of trading volume increases absolute returns by 0.000062 p.p., while the mean illiquidity is 0.12. In comparison, the annual average of Amihud ratio of NYSE stock from 1963 - 1996 is 0.000000337 (Amihud (2002), table 1). Market illiquidity exhibit less variation, but is overall illiquid, with mean of 0.12,

median of 0.13, and ranges from 0.0002 to 0.53. The token mean-adjusted illiquidity shows that while most tokens are more liquid than the market, as the 75th percentile is 0.02, a token can be 300 times more illiquid than the market.

The log market cap ranges from 3.77 to 23.30, or $\$43$ to $\$13.15$ billion, with mean of $\$2.6$ million. Beta, a token's quarterly co-movement with the market, ranges from -1797 to 6.44, with mean 0.68. Quarterly standard deviation of token returns ranges from 0 to 250.9, with mean and median 13.6 and 9.7, respectively. In contrast, the mean and median annual standard deviation of returns for the NYSE are 2.08 and 2.07, respectively.

For the onchain market maker variables, on average, 54% and 52% of all tokens received and sent, respectively, are by onchain market makers. Average net purchase of a token by market makers is 2.3% of all tokens transferred. Lastly, on average 19.3% of onchain market makers are whales.

For token holding behavior, on average, 57.5% of a token's addresses have positive balance, while 39.8% no longer hold that token. [6] Lastly, new addresses account for only 0.1% of a token's addresses on average.

[6]Negative-balance addresses are possible. The addresses from which tokens are minted have negative balance because they only send, but not receive.

Table 3.1: Summary Statistics of Token Characteristics

The illiquidity measure, Token Amihud$_{i,w}$, is week w's average of the daily ratio of absolute return to the dollar volume of token i. Excluded are tokens with extreme top and bottom 0.5% of illiquidity each day. The Market Amihud$_{i,w}$ is the simple average of Token Amihud$_{i,w}$ across all tokens each week. The mean-adjusted illiquidity AmihudMA$_{i,w}$ is Token Amihud$_{i,w}$ over Market Amihud$_{i,w}$. Returns$_{i,w}$ is percent change in i's price between the last days of w and $w-1$. Size$_{i,w}$ is the weekly average of i's daily market capitalization. Beta$_{i,q}$ is estimated each quarter from $R_{i,q} = \alpha_{i,q} + \text{beta}_{i,q}\text{RM}_q + \epsilon_{i,q}$. SDRET$_{i,q}$ is the quarterly standard deviation of daily returns. Buy and sell shares$_{i,w}$ are weekly averages of the daily ratio of the amount of i received or sent by onchain market makers over total transferred. Net buy share$_{i,w}$ is buy share subtract sell share. Whales$_{i,w}$ is the average of the daily percent of OMMs who are whales. Positive and Zero Address shares$_{i,w}$ are the average daily addresses with positive balance and zero balance, respectively, divided by the total number of addresses for i. New address share$_{i,w}$ is the weekly average of addresses that are new divided by total addresses. Weeks with fewer than 100 tokens are excluded. Data include 892 ERC-20 tokens over 164 weeks.

	N	mean	std	min	25%	50%	75%	max
Returns	75726	-0.579969	32.852758	-694.774410	-12.591901	-0.779794	10.290876	884.454220
Lag Returns	74827	-0.663473	32.682548	-694.774410	-12.689872	-0.846566	10.227972	884.454220
Token Amihud	74012	0.121542	0.894170	0.000000	0.000006	0.000062	0.001448	45.809818
Market Amihud	75726	0.122241	0.083824	0.000224	0.051349	0.127071	0.170733	0.528845
AmihudMA	74012	1.000000	6.179278	0.000000	0.000085	0.000962	0.020385	300.711880
Lag AmihudMA	73153	0.998269	6.172498	0.000000	0.000085	0.000969	0.020430	300.711880
Log Size	75726	14.767079	2.188704	3.774402	13.277742	14.767999	16.208253	23.304623
Lag Log Size	74812	14.759834	2.185219	3.774402	13.274539	14.764498	16.198299	23.304623
Beta	75726	0.684780	6.570113	-1797.216700	0.370855	0.672932	1.076485	6.438610
Lag Beta	75726	0.616270	13.314220	-1797.216700	0.381807	0.720340	1.161042	26.997532
SDRET	75726	13.601522	13.821433	0.000000	6.956965	9.725159	15.047136	250.900680
Lag SDRET	75709	13.113120	12.632989	0.000000	6.891344	9.627502	14.713506	250.900680
Buy Share	65004	0.541938	0.173705	0.000000	0.439443	0.539748	0.660821	1.000000
Sell Share	65004	0.519102	0.170944	0.000000	0.420059	0.514120	0.632611	1.000000
Net Buy Share	65002	0.022898	0.140404	-1.000000	-0.022807	0.005029	0.079029	1.000000
Whales	65004	19.345641	19.873626	0.000000	0.000000	16.269842	29.369113	100.000000
Positive Address Share	75726	0.575012	0.217723	0.000222	0.399245	0.559631	0.751246	0.999637
Zero Address Share	75726	0.397741	0.208619	0.000000	0.227967	0.409377	0.559823	0.999554
New Address Share	75726	0.001206	0.006006	0.000000	0.000047	0.000186	0.000672	0.306161

Table 3.2: Correlations

Pairwise correlations between regression variables. ** indicates statistical significance at 5%.

	Lag Beta	Lag SDRET	Buy Share	Sell Share	Net Buy Share	Positive Address Share	Zero Address Share	New Address Share	Returns	Lag Amihud MA	Lag Log Size	Whales
Lag Beta	1											
Lag SDRET	-0.153**	1										
Buy Share	-0.005	-0.038**	1									
Sell Share	0.004	-0.050**	0.668**	1								
Net Buy Share	-0.010**	0.013**	0.423**	-0.391**	1							
Positive Address Share	-0.014**	0.080**	-0.148**	-0.146**	-0.004	1						
Zero Address Share	0.014**	-0.087**	0.144**	0.141**	0.006	-0.979**	1					
New Address Share	-0.055**	0.026**	0.004	0.007	-0.004	0.003	-0.009**	1				
Returns	0.013**	-0.001	0.008**	0.031**	-0.028**	-0.008**	0.008**	0.050**	1			
Lag Amihud MA	0.001	0.227**	-0.017**	-0.005	-0.015**	-0.010**	0.002	-0.009**	0.008**	1		
Lag Log Size	0.007	-0.321**	0.124**	0.072**	0.065**	-0.137**	0.144**	0.114**	-0.048**	-0.202**	1	
Whales	0.004	-0.045**	0.216**	0.200**	0.024**	-0.107**	0.097**	0.034**	0.007	-0.025**	0.062**	1

Table 2 displays the pairwise correlations, where ** indicates statistical significance at 5% p-value. Return is positively correlated with buy and sell shares, meaning more OMM activity is associated with higher returns. However, the correlation coefficient for sell share is nearly four times that of buy share, which seems to indicate that OMM sell more when prices are high. OMMs' net purchase share is negatively correlated with returns, meaning that MM are net buyers when prices drop, that is, they are buying from everyone else during a sell-off.

Additionally, a higher fraction of new addresses is associated with higher returns, this could reflect either i) entities that have never transacted token i before started to do so (extensive margin) ii) entities who have transacted this token before created additional addresses to do so (intensive margin). The extensive margin could be a sign of herding behavior.

The lag of mean-adjusted Amihud ratio is positively correlated with returns, suggesting that liquidity premium might exist in crypto markets. Onchain market maker activity, as represented by buy share and sell share, is negatively correlated with illiquidity, meaning the OMMs trade more when when the market is more liquid. The fraction of OMM who are whales is also negatively correlated with illiquidity, perhaps because OMMs who own more tokens also trade more tokens, therefore providing more liquidity.

3.5.1 Regression Results

The model is estimated for 164 weeks, generating 164 sets of coefficients β_{jw}. The averages of the coefficients and t-statistics (in parentheses) across all 164 weeks are displayed in Table 3. Column (1) is the baseline model in the tradition of Fama and MacBeth (1973). Columns (2) - (5) add the onchain market maker variables, and columns (6) - (8) show results of token activity variables.

The negative and significant average coefficient for lag returns is evidence of mean reversion in all specifications. In the baseline model, a 1 p.p. increase in returns is followed by a 0.18 p.p. decrease the following week. Similar to equity, large tokens have lower returns, as a 1% increase in market cap decreases returns by 0.0032 p.p. The risk variables beta and return volatility are not significant across all specifications.

The coefficient on the mean-adjusted Amihud ratio shows significant liquidity premium for a few specifications. In the baseline regression, the average cross-section illiquidity coefficient is 0.088 with t-stat of 2.13, meaning if a token's illiquidity relative to the market increases by 1, which is 100 p.p., the following *week*'s non-annualized return increases by 0.088 p.p. In comparison, the average coefficient for stocks is 0.112 per *month* (non-annualized) in Amihud 2002. As the standard deviation of lag mean-adjusted Amihud is 6.17, a one standard deviation change in token illiquidity is associated with 0.54 p.p.increase in following week's return. When market makers are added, illiquidity's significance drops, with p-values between $5\% - 10\%$.

The share of tokens sent (sold) by market makers has significant coefficient of 3.59. As sell share ranges from 0 to 1, a 10 p.p. increase in tokens sent by market makers is correlated with 0.359 p.p.increase in weekly contemporaneous returns. This means that market makers tend to sell as price rises. On the other hand, the share of tokens bought by market makers (buy share) is insignificant. Market makers' net purchase of tokens as fraction of total transferred is negatively associated with returns, meaning market makers are net buyers when prices are low. This is consistent with the findings in the equity market (Comerton-Forde et al. (2010)). This shows that the OMMs are providing liquidity by temporarily absorbing the buy and sell pressure. A 10 p.p. increase in net purchase share is associated with a 0.622 p.p. decrease in returns.

Among the onchain activity variables, only new address share is significant. A 1 p.p. rise in new address share is associated with a 4.69 p.p. increase in return, which is economically significant. Note that the average new address share is only 0.1%, so a 1 p.p. change represents a large jump in new interest in a token and additional interest by existing holders.

Table 3.3: Cross Section Regression Results

$$R_{iw} = \alpha_w + \beta_{0w} R_{i,w-1} + \sum_{j=1}^{n} \beta_{jw} X_{ji,w-1} + \sum_{j=n+1}^{J} \beta_{jw} X_{ji,w} + \epsilon_{iw}$$

This table represents the means of coefficients from the weekly cross-sectional regression of stock return in percent on lagged financial variables ($X_{0i,w-1}...X_{ni,w-1}$) and contemporaneous onchain variables ($X_{n+1i,w}...X_{Ji,w}$). The data includes 164 weeks from 2018 to 2021, resulting in 164 sets of coefficients. T-statistics of the means are in parenthesis, * and ** denote statistical significance at 10% and 5% levels.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag Returns	-0.181**	-0.169**	-0.17**	-0.169**	-0.169**	-0.181**	-0.181**	-0.185**
	(-19.8)	(-18.794)	(-18.839)	(-18.488)	(-18.468)	(-19.697)	(-19.646)	(-20.246)
Lag AmihudMA	0.088**	0.105*	0.104*	0.097	0.102*	0.084**	0.084**	0.093**
	(2.129)	(1.741)	(1.729)	(1.629)	(1.709)	(2.027)	(2.02)	(2.255)
Lag Log Size	-0.317**	-0.305**	-0.344**	-0.253**	-0.308**	-0.328**	-0.325**	-0.453**
	(-2.776)	(-2.488)	(-2.809)	(-2.087)	(-2.548)	(-2.876)	(-2.833)	(-3.758)
Lag Beta	0.042	0.259	0.253	0.327	0.287	0.088	0.074	0.008
	(0.126)	(0.685)	(0.681)	(0.869)	(0.756)	(0.264)	(0.219)	(0.025)
Lag SDRET	-0.077	-0.071	-0.074	-0.071	-0.07	-0.074	-0.074	-0.09
	(-1.325)	(-1.099)	(-1.137)	(-1.102)	(-1.083)	(-1.271)	(-1.27)	(-1.538)
Buy Share		-0.257						
		(-0.302)						
Sell Share			3.585**					
			(4.139)					
Net Purchase Share				-6.222**				
				(-5.896)				
Whale Share					0.015**			
					(2.042)			
Zero Address Share						1.018*		
						(1.782)		
Positive Address Share							-1.033*	
							(-1.923)	
New Address Share								469.159**
								(4.831)
Constant	3.028	2.951	1.648	2.175	2.56	2.727	3.668	4.696*
	(1.285)	(1.158)	(0.656)	(0.867)	(1.046)	(1.153)	(1.549)	(1.939)

3.6 Conclusion

This chapter combines onchain and financial data to study the relationship between cryptocurrency returns, liquidity and onchain market maker activity. Cross section analysis shows the liquidity premium in cryptocurrency market is greater than that in equity market. Also consistent with studies of equity, onchain market makers' inventory is negatively correlated with returns. This indicates that OMMs provide liquidity by accommodating buying and selling pressure.

APPENDIX A

Blockchain Fundamentals

A.1 Cryptographic Primitives

A.1.1 Cryptographic Hash Functions

A cryptographic hash function has the additional properties of collision resistance, hiding, and puzzle friendliness (optional). A hash function H is collision resistant if it is infeasible to find two values, x and y, such that $x \neq y$ but $H(x) = H(y)$. Theoretically, collisions exist because the set of inputs is greater than the set of outputs; however, a secure cryptographic hash function makes the process of finding collisions so computationally intense that, in practice, collisions are infeasible to find. A hash function is said to be hiding if when a secret value r is chosen from a probability distribution with a large variance, given $H(r||x)$ (the hash of the concatenation of r and x), it is infeasible to find x. r is a randomly generated value that is used only once. A hash function is puzzle friendly if for every possible n-bit output value y, if a non-secret r is chosen from a spread out distribution, then it is infeasible to find x such that $H(r||x) = y$ in time significantly less than 2^n.

A.1.2 Digital Signatures

The second cryptographic primitive, digital signature, is an algorithm that satisfies two properties i) only the user can make the signature, but anyone can verify that it is valid, ii) the signature for a

particular message cannot be forged to other messages. A digital signature scheme consists of all three of the following algorithms:

- (sk, pk) :=generateKeys($keysize$) The generateKeys method takes a key size and generates a key pair. The secret key sk is private and used to sign messages. The public verification key pk is given to everybody, so anyone can verify your signature.
- sig :=sign(sk,$message$) The sign method takes a message and a secret key, sk, as input and outputs a signature for $message$ under sk.
- isValid: = verify(pk,$message$,sig). The verify method returns true if sig is a valid signature for $message$ under public pk

A.1.3 Transaction Propagation

If Alex wants to send Bob a Bitcoin, this transaction will be broadcasted to a few nodes in the Bitcoin network. Any node receiving the transaction first verifies the transaction's validity, then passes it along to the nodes that it is connected with. This way the transaction gets propagated to all of the nodes. How soon a node can receive a transaction depends on how many other nodes the transaction goes through to reach it. Because there is no global clock, a node only knows when it received each transaction, but not when each transaction was initially broadcasted. Therefore, it is impossible to determine the order in which the transactions actually occurred. If node A is selected to append the newest block (block n), it gets to decide both which transactions to include in that block and the order of these transactions.[1]

A.1.4 Bitcoin Transactions

Bitcoins are fungible, which means individual coins are not identifiable. A Bitcoin transaction is a message that consists of inputs that identify the previous transactions from which the coins

[1]A Bitcoin block is aaround 1MB in size, and can contain >2500 transactions of 300 - 400 bytes each. Bitcoin's highest daily average transaction per block occurred on March 26, 2019, for an average of 2734.4 transactions per block.

were received, and outputs that list the addresses to which coins are sent and the amount to be sent to each address. In this sense, a Bitcoin transaction is a redemption of previous transactions. For example, if address A sent six bitcoins to address B (transaction 1), but address B wants to send five of these coins to address C (transaction 2), then the input of transaction 2 would specify transaction 1, and the output of transaction 2 would specify five coins to address C and one coin back to address B itself. [2] A.1 shows the contents of an actual Bitcoin transaction.

Figure A.1: Contents of a Bitcoin Transaction

Metadata contains housekeeping information. The ”hash” field provides the hash of the entire transaction, which serves as the transaction ID. ”vin_sz” and ”vout_sz” are the number of inputs and outs. ”lock_time” specifies the earliest time the transaction can be added to the blockchain. *inputs* contains the hashes of previous transactions whose coins are being redeemed. ”scriptSig” are the signatures required to redeem the coins. *output* contains the address (public key) to which coins are being sent, and the quantity of coins to be sent.

[2]The Bitcoin blockchain does not contain the account balance of addresses. Verifying an address' account balance requires tracking all of the transactions that the address has been involved in, that is, calculating the sum of all coins that the address has ever received, and subtracting the sum of all coins that the address has ever sent. Instead, Bitcoin uses the unspent transaction output (UTXO) model. An unspent transaction output is a transaction output that can be used as input of another transaction. It is much simpler to track only the transactions through which the address received the coins to be spent.

APPENDIX B

Technical Appendix

B.1 Risk Neutral

B.1.1 Equity

$$\mathcal{L} = \lambda\beta(B_0 R) + (1-\lambda)\beta^2(B_0 R^2 + e[(y_1-\omega)R + y_2 - \omega]) + \mu(W - B_0 - eq) \tag{B.1}$$

$$\begin{aligned} \partial B_0: &\quad \lambda\beta R + (1-\lambda)\beta^2 R^2 = \gamma \\ \partial e: &\quad (1-\lambda)\beta^2[(y_1-\omega)R + y_2 - \omega] = \gamma q \\ \partial \mu: &\quad W = B_0 - eq \\ &\quad q = \frac{(1-\lambda)[(y_1-\omega)R + y_2 - \omega]}{R^2} \end{aligned} \tag{B.2}$$

Required rate of return: $R^E = \frac{(y_1-\omega)R + y_2 - \omega}{q} = \frac{R^2}{1-\lambda}$

$$\begin{aligned} c'_{\text{2equity}} &= (1-e)[(y_1-\omega)R + y_2 - \omega] \\ &= (y_1-\omega)R + y_2 - \omega - \frac{IR^2}{1-\lambda} \end{aligned} \tag{B.3}$$

B.1.2 Tokens

$$\mathcal{L} = \lambda\beta(B_0 R_f + \phi_1 T_0) + (1-\lambda)\beta^2[(B_0 R_f + \phi_1 T_0)R_f + \phi_2(1-\phi_1)T_0] + \mu(W - B_0 - T_0 p_0) \tag{B.4}$$

$$\begin{aligned}
\partial B_0: &\quad \lambda\beta R_f + (1-\lambda)\beta^2 R_f^2 = \mu \\
\partial T_0: &\quad \lambda\beta\phi_1 + (1-\lambda)\beta^2[\phi_1 R_f + \phi_2(1-\phi_1)] = \mu p_0 \\
\partial \mu: &\quad W = B_0 - T_0 p_0
\end{aligned}$$

$$p_0 = \frac{\phi_1}{R_f} + \frac{(1-\lambda)\phi_2(1-\phi_1)}{R_f^2} \tag{B.5}$$

Gross required rate of return (sell at $t = 1$ for 1):

$$\mathbb{E}_0(R^T) = \frac{1}{p_0} \tag{B.6}$$

$$\begin{aligned}
c'_{\text{2token}} &= T'_2 p_2 + B'_1 R_f - \omega \\
&= y_2 - \phi_2(1-\phi_1)T_0 + R_f(y_1 - \phi_1 T_0 - \omega) - \omega \\
&= (y_1 - \omega)R_f + y_2 - \omega - (\phi_2(1-\phi_1) + \phi_1 R_f)\frac{I}{p_0} \\
&= (y_1 - \omega)R_f + y_2 - \omega - (\phi_2(1-\phi_1) + \phi_1 R_f) I R^T \\
&= (y_1 - \omega)R_f + y_2 - \omega - I R_f^2 \left(\frac{\phi_2(1-\phi_1) + \phi_1 R_f}{\lambda R_f \phi_1 + (1-\lambda)(\phi_2(1-\phi_1) + \phi_1 R_f)}\right)
\end{aligned} \tag{B.7}$$

Result 1

$$\frac{\partial c'^{,T}_{2,n}}{\partial \phi_1} = I R^2 \frac{R\lambda\phi_2}{[\phi_1(R - (1-\lambda)\phi_2) + (1-\lambda)\phi_2]^2} \geq 0 \tag{B.8}$$

$$\frac{\partial c'^{,T}_{2,n}}{\partial \phi_2} = -I R^2 \frac{(1-\phi_1)\lambda\phi_1 R}{[\phi_1(R - (1-\lambda)\phi_2) + (1-\lambda)\phi_2]^2} \leq 0 \tag{B.9}$$

B.2 Constant Relative Risk-Averse Investors

B.2.1 Equity

$$\mathcal{L} = \lambda\beta\frac{(B_0 R_f)^{1-\sigma} - 1}{1-\sigma} + (1-\lambda)\beta^2\frac{(B_0 R_f^2 + e\Pi)^{1-\sigma} - 1}{1-\sigma} + \gamma(W - B_0 - eq)$$

$$
\begin{aligned}
\partial B_0: &\quad \lambda\beta(B_0R)^{-\sigma}R + (1-\lambda)\beta^2(B_0R^2+e\Pi)^{-\sigma} = \gamma \\
\partial e: &\quad (1-\lambda)\beta^2(B_0R^2+e\Pi)^{-\sigma}\Pi = \gamma q \\
\partial \gamma: &\quad W = B_0 + eq
\end{aligned}
$$

$$
\begin{aligned}
q_a &= \frac{(1-\lambda)\beta^2(B_0R+e\Pi)^{-\sigma}\Pi}{\lambda\beta(B_0R)^{-\sigma}R + (1-\lambda)\beta^2(B_0R+e\Pi)^{-\sigma}R^2} \\
&= \frac{(1-\lambda)\Pi}{R^2\left[\lambda(\frac{B_0R^2+e\Pi}{B_0R})^{\sigma}+1-\lambda\right]}
\end{aligned}
$$

$$
R_a^E = \frac{\Pi}{q} = \frac{R^2}{1-\lambda}\Big[\lambda(\frac{B_0R^2+e\Pi}{B_0R})^{\sigma}+1-\lambda\Big]
$$

$$
\begin{aligned}
c_{2,a}^{'\mathrm{E}} &= (1-e)\Pi = (1-\frac{I}{q})\Pi \\
&= \left[1 - \frac{IR^2\left[\lambda(\frac{B_0R^2+e\Pi}{B_0R})^{\sigma}+1-\lambda\right]}{(1-\lambda)\Pi}\right]\Pi \\
&= \Pi - \frac{IR^2}{1-\lambda}\Big[\lambda(\frac{B_0R^2+e\Pi}{B_0R})^{\sigma}+1-\lambda\Big]
\end{aligned}
$$

B.2.2 Token

$$
\begin{aligned}
\mathcal{L} = &\lambda\beta\frac{(B_0R_f+\phi_1T_0)^{1-\sigma}-1}{1-\sigma} \\
&+ (1-\lambda)\beta^2\frac{[(B_0R_f+\phi_1T_0)R_f+\phi_2(1-\phi_1)T_0]^{1-\sigma}-1}{1-\sigma} + \mu(W-B_0-T_0p_0)
\end{aligned}
$$

$$
\begin{aligned}
\partial B_0: &\quad \lambda\beta(B_0R+\phi_1T_0)^{-\sigma}R + (1-\lambda)\beta^2[(B_0R+\phi_1T_0)R+\phi_2(\phi_1)T_0]^{-\sigma}R^2 = \mu \\
\partial T_0: &\quad \lambda\beta(B_0R+\phi_1T_0)^{-\sigma}\phi_1 + (1-\lambda)\beta^2[(B_0R+\phi_1T_0)R+\phi_2(\phi_1)T_0]^{-\sigma}[\phi_1R+\phi_2(1-\phi_1)] \\
&= \mu p_0 \\
\partial \mu: &\quad W = B_0 + T_0p_0
\end{aligned}
$$

Let $\Phi = \phi_1 R + \phi_2(1-\phi_1)$.

$$
\begin{aligned}
p_{0,a} &= \frac{\lambda\beta(B_0R+\phi_1T_0)^{-\sigma}\phi_1 + (1-\lambda)\beta^2[(B_0R+\phi_1T_0)R+\phi_2(1-\phi_1)T_0]^{-\sigma}\Phi}{\lambda\beta(B_0R+\phi_1T_0)^{-\sigma}R + (1-\lambda)\beta^2[(B_0R+\phi_1T_0)R+\phi_2(1-\phi_1)T_0]^{-\sigma}R^2} \\
&= \frac{\frac{\lambda\beta\phi_1}{c_1^\sigma} + \frac{(1-\lambda)\beta^2[\phi_1R+\phi_2(1-\phi_1)]}{c_2^\sigma}}{\frac{\lambda\beta R}{c_1^\sigma} + \frac{(1-\lambda)\beta^2R^2}{c_2^\sigma}} \\
&= \frac{\beta\phi_1(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma) + (1-\lambda)\beta^2\phi_2(1-\phi_1)c_1^\sigma}{\lambda c_2^\sigma + (1-\lambda)c_1^\sigma} \\
&= \frac{\phi_1}{R} + \frac{(1-\lambda)\phi_2(1-\phi_1)}{R^2}\frac{c_1^\sigma}{\lambda c_2^\sigma + (1-\lambda)c_1^\sigma}
\end{aligned}
$$

$$
R_a^T = \frac{1}{p_{0,a}} = \frac{R^2(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma)}{\phi_1 R[\lambda c_2^\sigma + (1-\lambda)c_1^\sigma] + (1-\lambda)\phi_2(1-\phi_1)c_1^\sigma}
$$

$$
\begin{aligned}
c_{2,a}^{'\mathrm{T}} &= \Pi - (\phi_2(1-\phi_1)+\phi_1R)\frac{I}{p_0} \\
&= \Pi - (\phi_2(1-\phi_1)+\phi_1R)\frac{IR^2(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma)}{\phi_1R(\lambda c_2^\sigma + (1-\lambda)c_1^\sigma) + (1-\lambda)\phi_2(1-\phi_1)c_1^\sigma}
\end{aligned}
$$

BIBLIOGRAPHY

Acharya, Viral and Lasse H. Pedersen (2005) "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77 (2), 375–410.

Amihud, Yakov (2002) "Illiquidity and Stock Returns: Cross-sections and Time-series Effects," *Journal of Financial Markets*, 5 (1), 31–56.

Bakos, Yannis and Hanna Halaburda (2019) "The Role of Cryptographic Tokens and ICOs in Fostering Platform Adoption," *New York University Working Paper*.

Bancor (2017) "Bancor Network Token (BNT) Contribution Token Allocation Terms."

Brasher, Elizabeth, Jason L. Kropp, and David Westernberg (2020) "2020 Venture Capital Report."

Catalini, Christian and Joshua S Gans (2019) "Some Simple Economics of the Blockchain," *Rotman School of Management Working Paper No. 2874598*.

Chod, Jiri and Evgeny Lyandres (2021) "A Theory of ICOs: Diversification, Agency, and Information Asymmetry," *Management Science, forthcoming*.

Comerton-Forde, Carole, Terrence Hendershott, Charles Jones, Pamela Moulton, and Mark Seasholes (2010) "Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues," *Journal of Finance*, 65 (1), 295–331.

Cong, Lin William, Xi Li, Ke Tang, and Yang Yang (2021a) "Crypto Wash Trading," *Working Paper*.

Cong, Lin William, Ye Li, and Neng Wang (2021b) "Tokenomics: Dynamic Adoption and Valuation," *Review of Financial Studies*, 34 (3), 1105–1155.

Diamond, Douglas W. and Philip H. Dybvig (1983) "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy*, 91 (3), 401–419.

Etherscan (2021) "Ethereum Network Transaction Fee," *Etherscan.io*.

Fahlenbrach, R. and M Frattaroli (2021) "ICO investors," *Financial Market Portfolio Management*, 35, 1–59.

Fama, Eugene and James MacBeth (1973) "Risk, return, and equilibrium: Empirical tests," *Journal of Political Economy*, 81, 607 – 636.

Feld, Brad and Jason Mendelson (2019) *Venture Deals: Be Smarter Than Your Lawyer And Venture Capitalist 4th Edition*: Wiley.

Foundation, Ethereum (2023) "Learn about Ethereum," *ethereum.org*.

Gan, Jingxing, Gerry Tsoukalas, and Serguei Netessine (2020) "Initial Coin Offerings, Speculation, and Asset Tokenization," *Management Scienc*, 67 (w), 914–931.

Goldstein, Itay, Deeksha Gupta, and Ruslan Sverchkov (2022) "Initial Coin Offerings As a Commitment to Competition," *Working paper*.

Gorman, Michael and William A. Sahlman (1989) "What do venture capitalists do?," *Journal of Business Venturing*, 4 (4), 231–248.

Gornall, Will and Ilya A Strebulaev (2021) "The Economic Impact of Venture Capital: Evidence from Public Companies," *Working paper*.

Gryglewicz, Sebastian, Simon Mayer, and Erwan Morellec (2021) "Optimal Financing with Tokens," *Journal of Financial Economics, forthcoming*.

Hendershott, Terrence and Mark Seasholes (2007) "Market Maker Inventories and Stock Prices," *American Economic Review PP*, 97 (2), 210–214.

Higgins, Stan, Alex Sunnarborg, and Pete Rizzo (2017) "$150 Million: Tim Draper-Backed Bancor Completes Largest-Ever ICO."

Hinman, William (2018) "Digital Asset Transactions: When Howey Met Gary (Plastic)."

Howell, Sabrina T, Marina Niessner, and David Yermack (2020) "Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales," *The Review of Financial Studies*, 33 (9), 3925–3974.

Hsu, David (2004) "What Do Entrepreneurs Pay for Venture Capital Affiliation?," *Journal of Finance*, 59 (4), 1805–1844.

Jacklin, Charles (1987) *Demand Deposits, Trading Restrictions, and Risk Sharing*: University of Minnesota Press.

Jayachandran, Praveen (2017) "The difference between public and private blockchain," *IBM Blog*.

Kaur, Avinash, Anand Nayyar, and Parminder Singh (2020) "Blockchain: a Path to the Future," *ACM Transactions on Programming Languages and Systems*.

Lamport, Leslie, Robert Shostak, and Marshall Pease (1982) "The Byzantine Generals Problem," *ACM Transactions on Programming Languages and Systems*, 4 (3), 382–401.

Li, Jiasun and William Mann (2020) "Digital Tokens and Platform Building," *Revise and Resubmit at Review of Financial Studies*.

Lin, Hai., Junbo Wang, and Chunchi Wu (2011) "Liquidity Risk and Expected Corporate Bond Returns," *Journal of Financial Economics*, 99, 628–650.

Liu, Yukun, Aleh Tsyvinski, and Xi Wu (2022) "Common Risk Factors in Cryptocurrency," *Journal of Finance*, 77 (2), 1133–1177.

Lyandres, Evgeny, Berardino Palazzo, and Daniel Rabetti (2020) "ICO Success and Post-ICO Performance," *Working paper*.

Malinova, Katya and Andreas Park (2018) "Tokenomics: When Tokens Beat Equity," *Working Paper*.

Mancini, L, A Ranaldo, and J Wrampelmeyer (2013) "Liquidity in the Foreign Exchange Market: Measurement, Commonality, and Risk Premiums," *The Journal of Finance*, 68, 1805–1841.

Narayanan, Arvind, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder (2016) *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*: Princeton University Press.

Pozzi, Daniel (2019) "ICO Market 2018 vs 2017: Trends, Capitalization, Localization, Industries, Success Rate," *Cointelegraph*.

Puri, Manju and Rebecca Zarutskie (2012) "On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms," *Journal of Finance*, 67 (6), 2247–2293.

Rin, Marco Da, Thomas Hellmann, and Manju Puri (2013) "A Survey of Venture Capital Research," *Handbook of the Economics of Finance*, 2 (A), 573–648.

Ritter, Jay (2011) "Initial public offering, updated tables," *Unpublished working paper*.

Rowley, Jason (2019) "Q4 2018 Closes Out A Record Year For The Global VC Market," *Crunchbase*.

Sahlman, William A. (1990) "The structure and governance of venture-capital organizations," *Journal of Financial Economics*, 27 (2), 473–521.

Security and Exchange Commission (2019) "Framework for 'Investment Contract' Analysis of Digital Assets."

——— (2022) "Overview of Capital-Raising Exemptions."

Sockin, Michael and Xiong Wei (2021) "A Model of Cryptocurrencies," *Working Paper*.

Szabo, Nick (1997) "The Idea of Smart Contracts," *Satoshi Nakamoto Institute*.

Zhang, Wei and Li Yi (2023) "Liquidity Risk and Expected Cryptocurrency Returns," *International Journal of Finance and Economics*, 28, 472 – 492.

Zhao, Huabing (2018) "Hash Pointers and Data Structures," *medium.com*.