An Emperical Analysis of the Ethereum Blockchain

Steve Lee

2017

Abstract

Since the introduction of Bitcoin and the underlying blockchain technology, several alternative protocols have been created. This paper explores one of those alternatives, Ethereum, to examine the behavior of users and infrastructure providers. While more research is needed for robustness, I find tentative results that suggest: 1) infrastructure providers (called "miners") are currently able to operate at a profit, suggesting there is not a competitive equilibrium. Still, it would take a new miner approximately four months to breakeven from the fixed equipment cost; 2) users are willing to pay higher transaction fees, on average, in times of increased congestion, which is consistent with existing literature on queuing theory; and 3) decreasing rewards for infrastructure providers correlate with an increased level of infrastructure. Possible explanations for this include a time lag between deciding to become a miner and actually obtaining the necessary equipment.

1 Introduction

Since 2008, cryptocurrencies and their underlying blockchain technology have captured the imagination of many. While the ecosystem is still fledgling, money is pouring in. As of December 9, 2017, Bitcoin and Ethereum together (the two largest public blockchains) account for nearly $300 billion worth of investment[1]. With this in mind, I try to answer some basic questions about the economics of the Ethereum network. First, is the infrastructure maintained in a competitive equilibrium, or are there unobserved entry costs that allow for infrastructure providers to make a profit? Second, if transactions are processed as in a congested queuing game, theory dictates that transaction fees will increase with congestion (Huberman, Leshno, and Moallemi, 2017). Can this relationship be confirmed empirically? Finally, do the infrastructure providers respond rationally to changes in the incentive structure?

To answer these questions, I exploit two protocol level differences between Bitcoin and Ethereum. The first difference involves Ethereum's gradual (and

---

[1] https://coinmarketcap.com/

at times unpredictable) decrease in the rewards paid to infrastructure providers (called "miners") for successfully adding a new block to the chain[2]. Second, the current proof of work algorithm allows miners to receive expected rewards that are directly proportional to the miners percent of total infrastructure. This is to say that a miner who contributes 10% of the total infrastructure can expect to receive 10% of the rewards.

2 Blockchains

Introduced under the pseudonym Satoshi Nakamoto, blockchains were proposed as a solution to the so-called "double spending" problem (Nakamoto, 2008), which, simply put, is the risk that a single unit of digital currency could be sent to two different recipients [3]. While Bitcoin was the first to use this new technology, other blockchains have since emerged with redefined protocols. One such version, Ethereum, was proposed in 2013 by Vitalik Buterin with the goal of extending the blockchain architecture to include a "Turing-complete" programming language. This modification allows the Ethereum blockchain to process any logic of arbitrary complexity and provides the basis for "smart contracts" that can execute any program stored in the chain. In this sense, Ethereum can be thought of as a global computer with a publicly visible ledger of assets that can be securely traded, rented, and borrowed either directly or through more complex conditional statements [4].

2.1 New Blocks

As described in Nakamoto (2008), the technical steps necessary to add a new block to the Bitcoin blockchain are as follows:

1. Broadcast new transactions to all nodes (called "miners").
2. Each node records the new transactions into a block.
3. Each node works on finding a difficult proof-of-work for its block.
4. When a node finds a proof-of-work, it broadcasts the block to all nodes.
5. Other nodes accept the block only if all transactions in it are valid and not already spent.

---

[2] The Ethereum Foundataion plans to (eventually) switch from a proof of work verification protocol to a proof of stake method. Thus, they have instrumented different methods of reducing the expected rewards paid to miners in order to encourage users to adopt the updated network.

[3] Traditionally, this problem has been solved by using trusted intermediaries (i.e. banks, PayPal, Venmo, and so on).

[4] For example, consider the condition where Ann will lend Bob access to her car for the day only if Bob sends $10 to Chris.

6. Nodes express their acceptance of the block by working on creating the next block in the chain, using the hash of the accepted block as the previous hash.

Each block in the chain contains information from the previous block, new transactions, and a proof-of-work identifier (called “nonce”) that helps ensure that previous blocks cannot be maliciously altered. Visually, we represent the simplest version of this process below:

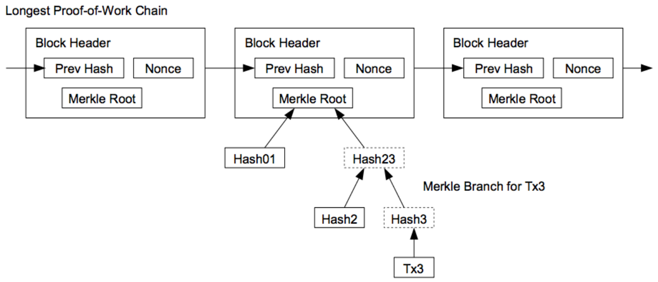

Figure 1: Elements of a simple block.

In addition to the computational aspect, incentives are used to encourage these nodes (again, called “miners”) to participate and behave properly. The current standard involves rewarding an accepted block’s creator with newly minted coins and all associated transaction fees.

While this represents the simplest case, the additional complexity in Ethereum (due to the Turing-complete programming language) introduces more variables for each block[5]. Further, even though both of these examples rely on proof of work to verify and secure each new block, alternative protocols are actively being explored with so-called “proof of stake” showing the most promise. Since the scope of this paper is to empirically analyze the existing Ethereum network, only proof of work will be explained and explored.

2.2 Proof of Work

The key innovation in Nakamoto (2008) that made blockchain architecture feasible was the “Proof of Work” algorithm. Proof of work is a piece of data that is hard to produce but easy to verify[6]. Conceptually, submitting a proof of work helps ensure that past blocks are infeasibly difficult to alter. At the technical level, proof of work involves searching for a value that, when cryptographically hashed [7] with the rest of the information in the block, returns a value that

---

[5] For a more detailed explanation of the Ethereum blockchain, see it’s technical paper at: https://ethereum.github.io/yellowpaper/paper.pdf.

[6] https://en.bitcoin.it/wiki/Proof_of_work

[7] Cryptographic hashes provide a unique random mapping from any input value of arbitrary size to a fixed-size alphanumeric value.

begins with some number of zeros determined by the difficulty. This step helps make revisions of prior blocks computationally infeasible[8]. A crude example of this step is shown below:

$$\begin{aligned} hash(previousHash + transactionRoot + wrongNonce) = ff686b0f5abb \\ hash(previousHash + transactionRoot + correctNonce) = 0000ccd546bc \end{aligned} \tag{1}$$

The latter is said to have "provided a valid proof of work". Thus, the goal for miners is to randomly check different values for the nonce until they find a value that, when concatenated with the rest of the relevant block information, yields a value with some leading number of zeros[9].

2.3 Mining Equipment

Each full node of the blockchain network is referred to as a "miner" and each miner expends it's computational resources to try and produce the next valid block in order to receive a reward[10]. To become a miner on the Ethereum network, one only needs access to the following equipment:

- Motherboard
- Graphics card (GPU)
- Hard drive
- Random access memory (RAM)
- Power supply
- Internet access

Once the "mining rig" is set up, some simple software installations allow the machine to start mining. Thus, by comparing the financial costs of mining to the expected rewards, we can derive an expression for the expected profit. Analytically we arrive at:

$$E(\pi)_n = h_n\left(\frac{R}{h} - \frac{P_e B_T}{e_h}\right) - F \tag{2}$$

---

[8] Consider, for example, a malicious attacker on the network. If they wanted to revise history, in the time it takes for all honest nodes to find a single new "nonce" value for the honest block, the attacker would have to compute a new "nonce" value for the malicious block they wanted to revise and every other block up the chain. Thus, as long as most of the nodes are honest, the network is secure.

[9] The exact number of leading zeros is determined by an automatically adjusting difficulty term. This difficulty is designed to produce blocks at a specific frequency. For Bitcoin, the difficulty produces blocks every 10 minutes, while for Ethereum the current block frequency is every 13 seconds.

[10] Unsurprisingly, the term miner refers back to real gold miners since digital miners "create" new cryptocurrency in the same way that gold miners "create" new gold.

Where,

$$
\begin{aligned}
R &= Expected\, reward,\ \frac{USD}{Block} \\
h_n &= Miner's\, hashrate\, (infrastructure),\ \frac{GH}{s} \\
h &= Total\, hashrate\, (infrastructure),\ \frac{GH}{s} \\
P_e &= Price\, of\, electricity,\ \frac{USD}{kW \cdot hr} \\
B_T &= Time\, per\, block,\ \frac{s}{Block} \\
e_h &= Efficiency\, of\, hashrate\, (infrastructure)\ \frac{GH}{kW \cdot hr} \\
F &= Fixed\, cost
\end{aligned}
\tag{3}
$$

To provide some perspective of the prices involved, a single new GPU will cost approximately \$300 and is only able provide about $\frac{1}{3{,}333{,}333}$ of the network's current infrastructure[11]. Thus, it is easy to see how even a relatively small scale mining rig could cost upwards of \$3,000. In the analysis section, we will do some back of the envelop calculations to determine if mining is currently profitable, and if so, how long it will take to recover your fixed equipment costs.

3 Queuing Theory

Queuing theory has a rich history in economic literature. In one example, Lui (1985) explores an equilibrium model for customers that can choose to pay a bribe in order to gain priority in a limited throughput queue. The basic conclusion is that as congestion increases, time sensitive customers will pay more to gain priority. Huberman, Leshno, and Moallemi (2017) apply these methods to the Bitcoin ecosystem and conclude that, in equilibrium, the system requires some minimum level of congestion in order to adequately fund the infrastructure providers.

While the state of Ethereum today is not in the equilibrium described by the aforementioned paper[12], it should still be the case that in times of increased congestion, time sensitive users are willing to pay more.

---

[11] One GPU can process about 30 MH/s while the entrie network currently processes more than 100,000,000 MH/s.

[12] The equilibrium model explored by Huberman, Leshno, and Moallemi (2017) analyzes system that is completely funded by transaction fees. This is to say that in their model, miners don't actually "mine" any new coins, but rather only collect transaction fees.

4 Data

Since all Ethereum blockchain data is publicly accessible, I retrieved and merged historical charts in order obtain a complete view, by day, of transaction fees, number of transactions, and other relevant indicators. A summary of the descriptive statistics is as follows:

Table 1: Descriptive statistics, by day, for the Ethereum network.

	units	mean	sd	min	max
totalTxs	–	110,402.1	147,790.6	0	791,746
exchangeRate	USD/ETH	78.8123	124.6133	0	475.24
aveTxFee	USD	0.0886927	.1593213	0	1.342702
blockReward	ETH	4.87921	.4767109	3	5
infrastructure	GH/s	22,945.8	35,196.59	23.7569	115,650.6
blockTime	s	16.14139	3.299734	4.46	30.31
N	861				

Here we see the wide variance across the independent variables. The exchange rate (denoted in USD per unit of Ethereum currency) was zero for the first few days of the network's history but recently has hit lifetime highs near $500 for one unit of "Ether". Likewise, total transactions and thus the average transaction fees were zero for a couple of the early days. Block reward refers to the expected, guaranteed payment to the winning miner (i.e. the miner who successfully adds a new block to the chain). The reward was initially set to 5 ETH per block of newly created "currency", but recently[13] they have reduced the reward to 3 ETH per new block[14].

The variable indicating infrastructure refers technically to the amount of computing power used in producing new blocks and is described in units of giga-hashes per second. Indeed, providing infrastructure to the network amounts to allocating computational "guessing" power towards finding and submitting a valid proof of work. There has been a dramatic increase in total infrastructure since the network went live in 2015, despite the recent decrease in expected reward. Finally, block time refers to the average time between blocks. This number can be targeted at the protocol layer by adjusting the difficulty required for a valid proof of work. Indeed, as the average time per block increases, the expected reward per unit time decreases for miners.

---

[13] October 16, 2017 the Ethereum Foundation updated the protocol.

[14] For more explanation of why, see the Ethereum Improvement Proposal number 186: https://github.com/ethereum/EIPs/issues/186

5 Analysis

In analyzing this data, I first test to see if increased congestion will in fact increase the average transaction fee. To do so I use the simple linear regression:

$$AveTxFee = \beta_0 + \beta_1 Txs + \beta_2 BlockTime + \beta_3 BlockSize + \epsilon \tag{4}$$

Note that the average transaction fee is described in USD and total transactions (Txs) is scaled to represent ten thousand transactions. Since congestion depends on the throughput of the network, I test the effects by restricting the dataset to days with some minimum number of transactions. The results are shown below:

Table 2: Effects of network congestion on the average transaction fee.

	All	10,000+ Txs	100,000+ Txs	300,000+ Txs
transactions	0.00747^{***}	0.0114^{***}	0.0139^{**}	0.0158^{**}
	(0.00175)	(0.00243)	(0.00515)	(0.00504)
blockTime	0.0149^{***}	0.0260^{***}	0.0361^{**}	0.0410^{***}
	(0.00278)	(0.00473)	(0.0109)	(0.0101)
blockSize	-0.000000819	-0.0000157^{*}	-0.0000347^{*}	-0.0000338^{*}
	(0.00000506)	(0.00000769)	(0.0000162)	(0.0000160)
N	854	716	224	116

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table shows that as congestion increases, users respond by paying higher transaction fees. More specifically, on days with more than 300,000 transactions, every additional 10,000 transactions increases the average transaction fee by almost two pennies. Further, we see that as block time increases, the average fee also increases. This makes sense as the cost of not getting included in a block will increase (for a time sensitive user) as block time increases.

Next, I analyze how infrastructure providrs (i.e. miners) respond to changing incentives. To do this I regress the total infrastructure on the ratio of reward to time. This ratio provides the expected reward per unit of time paid to winning miners. Thus, we would expect the coefficient on this number to be positive if miners are profit maximizing. The equation takes the form:

$$hashRate = \beta_0 + \beta_1 \frac{R}{T} + \beta_2 R + \beta_3 T + \beta_4 BlockSize + \beta_5 exchangeRate + \epsilon \tag{5}$$

Where R is the reward per block and T is the time between blocks. Regressing yields the following results:

	All	10,000+ GH/s	50,000+ GH/s
blockReward/Time	-7682.8	-941862.9***	-660257.9***
	(13208.2)	(45026.8)	(89164.0)
blockReward	-9998.7***	22674.2***	9981.5**
	(1883.3)	(2081.8)	(3667.6)
blockTime	192.1	-6490.5***	-4023.7***
	(369.2)	(398.2)	(748.1)
blockSize	3.730***	0.893**	0.843**
	(0.383)	(0.309)	(0.305)
exchangeRate	86.63***	5.202	-13.76
	(13.48)	(6.134)	(12.10)
N	861	286	169

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interestingly, I find that miners do not appear to respond rationally to changes in the incentive strucutre - or at least there is some unexplained relationship between the proliferation of new infrastructure and the expected reward per unit time. Next steps here would be to investigate the relationship in a time-lagged regression that accounts for the delay between ordering the necessary equipment and actually receiving it.

In considering the profitability of mining, we must return to the equation described in section 2.3 and provide some estimates:

$$E(\pi)_n = h_n\left(\frac{R}{h} - \frac{P_e B_T}{e_h}\right) - F \tag{6}$$

While exact global estimates are impossible due to fluctuation in energy prices and infrastructure efficiency, we can check for a range of reasonable values. In this spirit, I will use the following estimates based on the current environment:

$$R = \frac{3ETH}{Block}\frac{\$450}{ETH} = \frac{\$1350}{Block}$$

$$P_e = \frac{\$0.10}{kW \cdot hr}$$

$$B_T = \frac{13s}{Block} \tag{7}$$

$$e_h = \frac{583\,GH}{kW \cdot hr}$$

$$h = h_n = \frac{110,000\,GH}{s}$$

Here, the reward R, is found by multiplying the current reward per block (3 ETH per block) by the exchange rate ($450/ETH). The average time per block, B_T, is simply the rounded average of recent block times, and the price of electricity, P_e, is taken as the average price of electricity in the United States. The efficiency calculation is made by observing that the state of the art mining graphics card (GPU), the AMD Radeon RX 580, will produce about 30 $\frac{MH}{s}$ while using approximately 185 watts. While this omits the energy used by other components of the mining rig, for the purpose of estimation, I assume that the GPUs consume the vast majority of the power. Since there are constant returns to scale by individual hashing power, we can reduce the above equation to the following form and estimate the expected marginal returns to the entire network, $E(\pi)_M$, ignoring the fixed cost.

$$E(\pi)_M = \left(R - h\frac{P_e B_T}{e_h}\right) \tag{8}$$

I find that at current prices, mining yields an expected marginal profit to the network of $\frac{\$1,104}{Block}$. Note that this is only the operating profit, and recouping your fixed hardware costs could take considerable time at this rate[15]. Further, I find that given the prices and conditions listed above, as long as the exchange rate is greater than $80/ETH, mining will be operationally profitable.

6 Conclusion

In this paper I conducted an exploratory analysis of the Ethereum network to determine if users and infrastructure providers are acting rationally. I find

---

[15] Back of the envelope calculation suggests that the network is using at least $1 billion worth of hardware. At these prices, and realizing that there are approximately 6,500 new blocks added per day, it would take approximately 4 months to breakeven on your investment.

that, consistent with predictions based on existing queuing theory literature, average transaction fees increase as congestion increases in queues with limited throughput. Further, I find that infrastructure providers do not appear to respond as expected with respect to changes in the incentive structure. Finally, I estimated the profitability of acting as an infrastructure provider and conclude that while it is still operationally profitable mine on the Ethereum network, there is a large barrier to enter insofar as the upfront fixed cost of obtaining all the necessary equipment[16] is unlikely to be recovered. Thus, for someone who already has all the equipment, mining is still profitable, but factoring in the fixed entry cost will likely result in losses.

References

[1] Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System.* 2008

[2] Huberman, Leshno, and Moallemi. *Monopoly without a Monopolist: An Economic Analysis of the Bitcoin Payment System.* Working Paper, 2017.

[3] Lui. *An Equilibrium Queuing Model of Bribery.* Journal of Political Economy, 1985.

---

[16] This makes no attempt to estimate or include learning costs